```
TGCGTCACCTGCAGGCCCGGGCCGCGGGGTTGGTTTCCACCCTGGAGGTTGCTGACACCCTGTGCCCTCGGCTGACTTC

CAGCCGGTGGCACAGACGCCTCCAGGGGGCAGCACTCAAGCGCATCTTAGGAATGACAGAGTTGCGTCCCTCTCTGTTG

CCAGGCTGGAGTTCAGTGGCATGTTCTTAGCTCACTGAAGCCTCAAATTCCTGGGTTCAAGTGACCCTCCCACCTCAGC

                                                    M   K   I   K   A   E   K   N     8
CCCATGAGGACCTGGGACTACAGGACACAGCTAAATCCCTGACACGG ATG AAA ATT AAA GCA GAG AAA AAC    24

 E   G   P   S   R   S   W   W   Q   L   H   W   G   D   I   A   N   N   D   G     28
GAA GGT CCT TCC AGA AGC TGG TGG CAA CTT CAC TGG GGA GAT ATT GCA AAT AAC AGC GGG     84

 N   M   K   P   P   L   L   V   F   I   V   C   L   L   W   L   K   D   S   H     48
AAC ATG AAG CCG CCA CTC TTG GTG TTT ATT GTG TGT CTG CTG TGG TTG AAA GAC AGT CAC    144

 C   A   P   T   W   K   D   K   T   A   I   S   E   N   L   K   S   F   S   E     68
TGC GCA CCC ACT TGG AAG GAC AAA ACT GCT ATC AGT GAA AAC CTG AAG AGT TTT TCT GAG    204

 V   G   E   I   D   A   D   E   E   V   K   K   A   L   T   G   I   K   Q   M     88
GTG GGG GAG ATA GAT GCA GAT GAA GAG GTG AAG AAG GCT TTG ACT GGT ATT AAG CAA ATG    264

 K   I   M   M   E   R   K   E   K   E   H   T   N   L   M   S   T   L   K   K    108
AAA ATC ATG ATG GAA AGA AAA GAG AAG GAA CAC ACC AAT CTA ATG AGC ACC CTG AAG AAA    324

 C   R   E   E   K   Q   E   A   L   K   L   L   N   E   V   Q   E   H   L   E    128
TGC AGA GAA GAA AAG CAG GAG GCC CTG AAA CTT CTG AAT GAA GTT CAA GAA CAT CTG GAG    384

 E   E   E   R   L   C   R   E   S   L   A   D   S   W   G   E   C   R   S   C    148
GAA GAA GAA AGG CTA TGC CGG GAG TCT TTG GCA GAT TCC TGG GGT GAA TGC AGG TCT TGC    444
```

FIG. 1

```
 L   E   N   N   C   M   R   I   Y   T   T   C   Q   P   S   W   S   S   V   K   168
CTG GAA AAT AAC TGC ATG AGA ATT TAT ACA ACC TGC CAA CCT AGC TGG TCC TCT GTG AAA  504

 N   K   I   E   R   F   F   R   K   I   Y   Q   F   L   F   P   F   H   E   D   188
AAT AAG ATT GAA CGG TTT TTC AGG AAG ATA TAT CAA TTT CTA TTT CCT TTC CAT GAA GAT  564

 N   E   K   D   L   P   I   S   E   K   L   I   E   E   D   A   Q   L   T   Q   208
AAT GAA AAA GAT CTC CCC ATC AGT GAA AAG CTG ATT GAG GAA GAT GCA CAA TTG ACC CAA  624

 M   E   D   V   F   S   Q   L   T   V   D   V   N   S   L   F   N   R   S   F   228
ATG GAG GAT GTG TTC AGC CAG TTG ACT GTG GAT GTG AAT TCT CTC TTT AAC AGG AGT TTT  684

 N   V   F   R   Q   M   Q   Q   E   P   D   Q   T   F   Q   S   H   F   I   S   248
AAC GTC TTC AGA CAG ATG CAG CAA GAG TTT GAC CAG ACT TTT CAA TCA CAT TTC ATA TCA  744

 D   T   D   L   T   E   P   Y   F   F   P   A   F   S   K   E   P   M   T   K   268
GAT ACA GAC CTA ACT GAG CCT TAC TTT TTT CCA GCT TTC TCT AAA GAG CCG ATG ACA AAA  804

 A   D   L   E   Q   C   W   D   I   P   N   F   F   Q   L   F   C   N   F   S   288
GCA GAT CTT GAG CAA TGT TGG GAC ATT CCC AAC TTC TTC CAG CTG TTT TGT AAT TTC AGT  864

 V   S   I   Y   E   S   V   S   E   T   I   T   K   M   L   K   A   I   E   D   308
GTC TCT ATT TAT GAA AGT GTC AGT GAA ACA ATT ACT AAG ATG CTG AAG GCA ATA GAA GAT  924

 L   P   K   Q   D   K   A   P   D   H   G   G   L   I   S   K   M   L   P   G   328
TTA CCA AAA CAA GAC AAA GCT CCT GAC CAC GGA GGC CTG ATT TCA AAG ATG TTA CCT GGG  984

 Q   D   R   G   L   C   G   E   L   D   Q   N   L   S   R   C   F   K   F   H   348
CAG GAC AGA GGA CTG TGT GGG GAA CTT GAC CAG AAT TTG TCA AGA TGT TTC AAA TTT CAT 1044
```

FIG.1B

```
 E  K  C  Q  K  C  Q  A  H  L  S  E  D  C  P  D  V  P  A  L  368
GAA AAA TGC CAA AAA TGT CAG GCT CAC CTA TCT GAA GAC TGT CCT GAT GTA CCT GCT CTG 1104

 H  T  E  L  D  E  A  I  R  L  V  N  V  S  N  Q  Q  Y  G  Q  388
CAC ACA GAA TTA GAC GAG GCG ATC AGG TTG GTC AAT GTA TCC AAT CAG CAG TAT GGC CAG 1164

 I  L  Q  M  T  R  K  H  L  E  D  T  A  Y  L  V  E  K  M  R  408
ATT CTC CAG ATG AGG CGG AAG CAC TTG GAG GAC ACC GCC TAT CTG GTG GAG AAG ATG AGA 1224

 G  Q  F  G  W  V  S  E  L  A  N  Q  A  P  E  T  E  I  I  F  428
GGG CAA TTT GGC TGG GTG TCT GAA CTG GCA AAC CAG GCC CCA GAA ACA GAG ATC ATC TTT 1284

 N  S  I  Q  V  V  P  R  I  H  E  G  N  I  S  K  Q  D  E  T  448
AAT TCA ATA CAG GTA GTT CCA AGG ATT CAT GAA GGA AAT ATT TCC AAA CAA GAT GAA ACA 1344

 M  M  T  D  L  S  I  L  P  S  S  N  F  T  L  K  I  P  L  E  468
ATG ATG ACA GAC TTA AGC ATT CTG CCT TCC TCT AAT TTC ACA CTG AAG ATC CCT CTT GAA 1404

 E  S  A  E  S  S  N  F  I  G  Y  V  V  A  K  A  L  Q  H  F  488
GAA AGT GCT GAG AGT TCT AAC TTC ATT GGC TAC GTA GTG GCA AAA GCT CTA CAG CAT TTT 1464

 K  E  H  F  K  T  W  *                                        496
AAG GAA CAT TTT AAA ACC TGG TAA                                1468

GAAGATCTAATGCATCCTATATCCAGTAAGTAGAATTATCTCTTCATCTGGGACCTGGAAATCCTGAAATAAAAAAGGA

TAATGCAATAAACACAGTTGCAGGAAAGTATGTTAGCTATATACTATGAAGTACTCTTAGTTTACTTATGTTGAATGGC

TTAGCTATTAATACTCAAATTGAGTTAAAATGAAAATTCCTCCTTAAAAAATCAAACGTAATATGTATTACATTTCATG

GTACATTAGTAGTTCTTTGTATATTGAATAAATACTAAATCACCTA
```

FIG. 1C

```
TGCGTCACCTGCAGGCCCGGGCCGCGGGGTTGGTTTCCACCCTGGAGGTTGCTGACACCCTGTGCCCTCGGCTGACTTC

CAGCCGGTGGCACAGACGCCTCCAGGGGGCAGCACTCAAGCGCATCTTAGGAATGACAGAGTTGCGTCCCTCTCGGTTG

CCAGGCTGGAGTTCAGTGGCATGTTCATAGCTCACTGAAGCCTCAAATTCCTGGGTTCAAGTGACCCTCCTACCTCAGC

     M   R   T   W   D   Y   S   N   S   G   N   M   K   P   P   L   L   V   F    19
CCC ATG AGG ACC TGG GAC TAC AGT AAC AGC GGG AAC ATG AAG CCG CCA CTC TTG GTG TTT   57

 I   V   C   L   L   W   L   K   D   S   H   C   A   P   T   W   K   D   K   T    39
ATT GTG TGT CTG CTG TGG TTG AAA GAC AGT CAC TGC GCA CCC ACT TGG AAG GAC AAA ACT  117

 A   I   S   E   N   L   K   S   F   S   E   V   G   E   I   D   A   D   E   E    59
GCT ATC AGT GAA AAC CTG AAG AGT TTT TCT GAG GTG GGG GAG ATA GAT GCA GAT GAA GAG  177

 V   K   K   A   L   T   G   I   K   Q   M   K   I   M   M   E   R   K   E   K    79
GTG AAG AAG GCT TTG ACT GGT ATT AAG CAA ATG AAA ATC ATG ATG GAA AGA AAA GAG AAG  237

 E   H   T   N   L   M   S   T   L   K   K   C   R   E   E   K   Q   E   A   L    99
GAA CAC ACC AAT CTA ATG AGC ACC CTG AAG AAA TGC AGA GAA GAA AAG CAG GAG GCC CTG  297

 K   L   L   N   E   V   Q   E   H   L   E   E   E   E   R   L   C   R   E   S   119
AAA CTT CTG AAT GAA GTT CAA GAA CAT CTG GAG GAA GAA GAA AGG CTA TGC CGG GAG TCT  357

 L   A   D   S   W   G   E   C   R   S   C   L   E   N   N   C   M   R   I   Y   139
TTG GCA GAT TCC TGG GGT GAA TGC AGG TCT TGC CTG GAA AAT AAC TGC ATG AGA ATT TAT  117

 T   T   C   Q   P   S   W   S   S   V   K   N   K   I   E   R   F   F   R   K   159
ACA ACC TGC CAA CCT AGC TGG TCC TCT GTG AAA AAT AAG ATT GAA CGG TTT TTC AGG AAG  477
```

FIG.2A

```
 I   Y   Q   F   L   F   P   F   H   E   D   N   E   K   D   L   P   I   S   E   179
ATA TAT CAA TTT CTA TTT CCT TTC CAT GAA GAT AAT GAA AAA GAT CTC CCC ATC AGT GAA  537

 K   L   I   E   E   D   A   Q   L   T   Q   M   E   D   V   F   S   Q   L   T   199
AAG CTC ATT GAG GAA GAT GCA CAA TTG ACC CAA ATG GAG GAT GTG TTC AGC CAG TTG ACT  597

 V   D   V   N   S   L   F   N   R   S   F   N   V   F   R   Q   M   Q   Q   E   219
GTG GAT GTG AAT TCT CTC TTT AAC AGG AGT TTT AAC GTC TTC AGA CAG ATG CAG CAA GAG  657

 F   D   Q   T   F   Q   S   M   F   I   S   D   T   D   L   T   E   P   Y   F   239
TTT GAC CAG ACT TTT CAA TCA CAT TTC ATA TCA GAT ACA GAC CTA ACT GAG CCT TAC TTT  717

 F   P   A   F   S   K   E   P   M   T   K   A   D   L   E   Q   C   W   D   I   259
TTT CCA GCT TTC TCT AAA GAG CCG ATG ACA AAA GCA GAT CTT GAG CAA TGT TGG GAC ATT  777

 P   N   F   F   Q   L   F   C   N   F   S   V   S   I   Y   E   S   V   S   E   279
CCC AAC TTC TTC CAG CTG TTT TGT AAT TTC AGT GTC TCT ATT TAT GAA AGT GTC AGT GAA  837

 T   I   T   K   M   L   K   A   I   E   D   L   P   K   Q   D   K   A   P   D   299
ACA ATT ACT AAG ATG CTG AAG GCA ATA GAA GAT TTA CCA AAA CAA GAC AAA GCT CCT GAC  897

 H   G   G   L   I   S   K   M   L   P   G   Q   D   R   G   L   C   G   E   L   319
CAC GGA GGC CTG ATT TCA AAG ATG TTA CCT GGG CAG GAC AGA GGA CTG TGT GGG GAA CTT  957

 D   Q   N   L   S   R   C   F   K   F   H   E   K   C   Q   K   C   Q   A   H   339
GAC CAG AAT TTG TCA AGA TGT TTC AAA TTT CAT GAA AAA TGC CAA AAA TGT CAG GCT CAC 1017

 L   S   E   D   C   P   D   V   P   A   L   H   T   E   L   D   E   A   I   R   359
CTA TCT GAA GAC TGT CCT GAT GTA CCT GCT CTG CAC ACA GAA TTA GAC GAG GCG ATC AGG 1077
```

FIG.2B

```
 L  V  N  V  S  N  Q  Q  Y  G  Q  I  L  Q  M  T  R  K  H  L  379
TTG GTC AAT GTA TCC AAT CAG CAG TAT GGC CAG ATT CTC CAG ATG ACC CGG AAG CAC TTG 1137

 E  D  T  A  Y  L  V  E  K  M  R  G  Q  F  G  W  V  S  E  L  399
GAG GAC ACC GCC TAT CTG GTG GAG AAG ATG AGA GGG CAA TTT GGC TGG GTG TCT GAA CTG 1197

 A  N  Q  A  P  E  T  E  I  I  F  N  S  I  Q  V  V  P  R  I  419
GCA AAC CAG GCC CCA GAA ACA GAG ATC ATC TTT AAT TAC ATA CAG GTA GTT CCA AGG ATT 1257

 H  E  G  N  I  S  K  Q  D  E  T  M  M  T  D  L  S  I  L  P  439
CAT GAA GGA AAT ATT TCC AAA CAA GAT GAA ACA ATG ATG ACA GAC TTA AGC ATT CTG CCT 1317

 S  S  N  F  T  L  K  I  P  L  E  E  S  A  E  S  S  N  F  I  459
TCC TCT AAT TTC ACA CTC AAG ATC CCT CTT GAA GAA AGT GCT GAG AGT TCT AAC TTC ATT 1377

 G  Y  V  V  A  K  A  L  Q  H  F  K  E  H  F  K  T  W  *        478
GGC TAC GTA GTG GCA AAA GCT CTA CAG CAT TTT AAG GAA CAT TTT AAA ACC TGG TAA     1434

GAAGATCTAATGCATCCTATATCCAGTAAGTAGAATTATCTCTTCATCTGGGACCTGGAAATCCTGAAATAAAAAAGGA

TAATGCAATAAACACAGTTGCAGGAAAGTATGTTAGCTATATACTATGAAGTACTCTTAGTTTACTTATGTTGAATGGC

TTAGCTATTAATACTCAAATTGAGTTAAAATGAAAATTCCTCCTTAAAAAATCAAACGTAATATGTATTACATTTCATG

GTACATTAGTAGTTCTTTGTATATTGAATAAATACTAAATCACCTA
```

FIG.2C

FIG. 3A

ACATTTTAAGCTACTTATAGTCCTTGGAAATAGCAACAAATATCTTAGTTATTGGACTATTATAACCTTAGTCATCTTATTACTGCTTG
ATTATGAGACACTCTCCCTGCTAATCCTTAGAACATCTTGGTTCTTGGTACTTGACTTTTAGCCCCTCTGACATATAGTTGATGTCAGA
GTGTCTGGCATTTCAGTAGTGCTCTATTTTACAAATCCCAGTAAACTGCTCCACTGTGGCTTGTTTATGTGTTAATACTGCTTGTTTTC
TGTTATAAATTATTTTTTGCTTTGGAGTAAGATATCATCATTTTGCATAGCTACAAATCTGAAGTTAAAGAAAATTTTAAAAATGTAAT
TGTGGGAAAATAACAAATAGATCTGCTGAGATGGAGGCTTTGACTAATGTTTTAATAACAGGCAACAAAACAAAGAGGCAGGATATTTT
GGTCACAACTAAACCTAAATTAAATCCTCATACAAAGCCCCATTAAGATAAATGCTCAAATTCTGGGAACATTTCACTTGCTTTGCCAG
CAATTTTACCCTTCAGAGGGTGTGGATCTAATCAGGGGAACAAACTACCCTGGGCTTAATTCTCATTAACAGGGACTAATTTGTCAAAG
CGGCAGTACTAGCTGAAGTGATGGGTATGGAAGCATTCACTGTGAGGATTTTGCTGAGGTGCCTGGCACAGGGTAGGGGAACTCACCCA
GGCTGCAAGATGCTAACAGTTCAGGTTCAAGGTCTTAGTGTGGACTAAGGTGCAGTCAGGATGGGAACAGGTGCAACTTGGGCCAACAT
CAGTATGAAGGGCCTGATCTGAGGGCAGGGGAAGGAGGGGGCATTCTGGGAAGCAAGAGTTCCTGGTATCCTGTTGACCAGAGTCTTGG
CCCAAGGATCAACGTATGAATTAAAGTAGAAATACCAGAAACAAAGAAAGTTGGCAGAAACTAGGAGAAGCAGAGTCTCAGCCAACTGG
ACTGGGCTCAGCCTTGGCTACTGGCCCGGCAGATGATAGAAGAGAAAACCAGGAACCCAGGCTGAAGCCCAGTGGTTGGGCTGGCCACA
CACCATGCATAGCCTTAAAGGGGTGGCCTAAGGGCATGGTCCGCTCCAAAAAAGGAAAGGGGGCCCCAGAATATTTCTGAATCCCACTC
ACTGCCAGGGAAGAACCTCTCAATTCACTCAATAGTGCATTCTCCTGCTTCTCAATAGGCTAATACTCTAGAGAATATGGGGACAAGGG
GAGGAGGGTCTAGTGGAACAGGTCTAAACTGGCGTTTGAATTTTAAGATAAGTTAATCATACATTGGCTGGGTCAGCCATGTCTCTTAG
TCTTTACAAAAGTAGAACACAAAAAAATTCAATGGAAATCTACAGACACCTATTTGCAGATGAGGAAACACGGCTATGAAGATTGGGAA
GATTGGGAAGAACTGGCCAGGTGTGGTGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCTGGTGGATCACTTGAGGTCAGGA
GTTGGAGACCAGCCTGGGCAACATAGTAAAACCCTGTCTCTACTCAAATTACAAAAATCAGCAGGGCGTTGTGGTGCCCACCTGTAATC
CCAGCTATGCAGGAGGCTGAGGCAGGACAATCACTTGAACCTGGTAGGCGGAGGTTGCAGTGAGCCAAAATCACGCCACTGTACTCCAG
CCTGGGTGACAGAGCAAGACTTTGTTTAAAAAAAAAAAAAAAAAGGGAAGAACTAAAAATGTAATTTTCAAGGGGCTATCACAAATGGT
CCCAATAAAGAGAAAGCAGGACTCATGTTTAAGAAACCCATGAGATGTGTATGGACCTCATGGAAGAGCTCTTGCTTTCTAATGATCTA
CGTAACAGATGAAAAGCAGAGCATAGGGCTAAGGATGAAAATACAACAGTAATAAGGTATTAATATATTATTAAGAAAGCTAATGCTCC
ACATAAGCAGAGGACATTAAAGGGACTTTTTTTTTCTTAAGGATATCTTAATGTTTTAAATGAGAAGACATAGAAAGGGATAGGTCCAAC
TCTTGGGATTGTTGCAGGTTGGTTTCCATCGGAAGCACTCTGAGTCTGAGATTTGTATGCAGAAAATTAATTTGAATGTGCTTTTCAGA
TCACCCAGGTGGGGGAGGGAGGAAACCAGGACTGGGCAGAGAGAGGCTGGGCTGTAACCAAGTCACAACAAAGGTGTCAGCTGGTCCCA
TGGTGAATTCTGGACCTAGGATGGCTGATCCCAAGGCATTCCAAACTGGGGCAAGGAAGTTGTGCTTTAAAACTTCTCATTGACTGTCA
GTCACTGGGCATGAGCAGTCCCCAGGAAGGGGGGATGACCTTGAGCAAGGTGGATGTCTTCAGCCAAGGGCAAYCACTGGGAAGGAGAA
CCCAGCTATGAACTGTCAGCTGCCAACACTCCCAGCATCTGAGAGGATGAGGGCTTCAATTCTAAGGGCAGGGGCTCCAAGGGCAGGGG
TACGGATGGTGGAATCTGGGCAGTACCTTGTGGCTTCCACTACAGTCCACCCCTTGCACCACTTAGTTCCACTGGCTTTTTTTTTTTT

TTCTTTTCTGAGACAGTCTCACTCTGTCACCCAGGCTGGAGTGCGGTGGCACGATCTCGGCTCGCTGCAACCTCCGCCTCCCAGGTTCA
AGCAATTCTTGAACCTCCTGAGTAGCTGGGACTACAGATGTGTGCCACCACACCCAGCTAATTTTTTTGTATTTTTAGTAGAGACGGGGT
TTTACCGTGTTAGCCAGATTGGTCTCGATCTCCTGACCTCATGATCCGCCTGCTTTGGCCTCCCAAAGTGCTGGGATTACAGGTGTGAG
CCACCGCACACAGCCAGATCCACTGGCTTCTATATAATTTCTGGGTGAAGCTAATTCAGGATTCTGATGGACCTGTCTTCCCGAGGGAA
ACTTGTAAAAGGAAAGTTAGAGGGACAAACTATAGCCCCTGCCACAGCAGCTGCTGTCGAGGACAAAAATGGTGCTCCTCATTTCCCTT
AACCACCTGACCTAGATTCCCCTAACCCTTAGTGGGCACCTCTGTGGATGGAAGTGGTGGCTCACYKGKKGGRWKRWYCMRRWYYCWYM
YCCCTGAGTGGTCTGAGCTCCCAGTTACCAGGCCCTTCTCAGGCTGTGGCTGTTGCACTTACCTCCCCAGCCATCCCCCACTTTTTTTT
CTTGAGACTGGGTCTTGCTCTGTCACCCAGGCTGAAATGCAGTGGCATAACCTCAGCTCACTGCAGCCTTGATCTCCCAAGCTCAAGCC
ATCTTCTCACCTCTGCCTCCCAAGTGGCTGGGACTACAGGCACATGCCACCATGCCCAGCTAATATTTTTTATTTTTTATTTTTTTGTA
GCAATGGGATTTTGCCATGTTTCCCAGGCTGGGCTTGAACTCCTAAGCTCAAGCTATCCTCCCACCTCTGCTTCCCAAAGTGCTGGGAT
TACAGGCTTGAGTCACTGCATCTGGCCACATTTATTCCTTTTAAACGTTAAAATTGAATGCAGGATCACTGAGAGACAGGTGAGTGATT
ACCAGGGTGCCAAACATACCCTTCTCCTCCTTTCCTGCAGCTCTACCTCCTCCTGATGATCAGGACAATCATGTATGATGACTCCTTTC
CTTGACTGCTGCTCTCTCAGAAGGAACCCATTGTGTTGGGTGAGAACACATCATTTGAAATTTAGTAAGACTCTTGCTGTGCCTATGGT
AGAAGCATTCCCTCTCTGGGGCCAAGATCTTTAAATGCACAGAGTCCAAAGTCGTGGGAACCAAAGCAGAAATTAAAAAGGAGATGACT
GGGATTATGGTAAGAACTGTTTCCACCCTTGATTTGCTGCACCCATGTGTTCTACCTAGGAGATAGCACACCATATACTGGTTATTCAT
TTGGATTACATGCTGCATCCCGGAGAATGGGCACTGCATTCTCACTGGTCATCATGTCAGAGCCTGCGCTGCAGAGGCTTTCCCATTGC
TCTGTCAGTGTGTTATAGGGTCAGTGGATTTCATGGTCATGTGCCCACTGCTGCACCTCCATTCTTGTAAAATGGGTCCTCTGGTTCAA
TGTGATGCCATGTGGGATCTTGTGTCAATAGAATAAATACTCAGATGTTCTGGCTGAAGCTTTACAAGCAGAAAAGGCCAACCGATGAC
TGAAATAAGCGTTGAGCCCAGTCAAGATGAGTTCCTGCTCTTTCCAGGATAGACGGAGTCTAGTGTAGATCACTTGACATCAAGAGACT
GGCTGGTCTCCTTGAGGGATGGTGCTGTTCTGCATTCATCATCCTTGATGAATGAGGGACCCTGCTATTGGGCTCATGTACAGCCCCCA
TCTCTGCCACAATGAGCGCTCCATTCATGTTCCTATTGTGCCAACACTAGGGTGTCTGTAATCACTGAAAACATTATTGCTATCATTAT
TATTATTTTTTTTTTTTGAGACAGAGTCTCGCTCTGTCGCCAAGGCTGGAGTGCAGTGGCACGATCTCAGCTCACTGCAACCTCTGCCT
CCCGGCTTCAAGTGATTCTCCCGCCTCAGCCTCCAGAGTAGCTGGGATTATAGGCATGCGCCACCACGCCTGGCTAATTTTTGTATTTT
TAGTAGAGACAGTCTTTTGCCATATTAGTCTGTCTGGTCTCGAACTCCTGACCTCAGGTGATCTGCCCGCCTTGGCCTTCCGGAGTGCT
AGGATTATAGGCGTGAGCCACCACTTGCTATTATTATGTTGAGAAAACTGTTTTCAATTATAAATAAGAAAAAATAAAAGATTATATTT
TGCCTTTATTCCTTCTCTAATGCTGTTCTTTAAGTAGATGTGAATTTCTGAACTACATACTTTTTCTTTACTCTTGAGAGGTTGTTTGG
AGGTTCCAGCAGGGGACCACAGCTACTCGTATACCCTTGACCAAAGACTGGTCCTTGTCTATCAAGGATGGTCGTCTTCTTCCACCAAG
CACACAGCTTCTGGAGGGACGCACATGGAGTGGTGAGGGAGGAAGGGGACACCCGCCTAGCCAGCTAGATCAGCCAAGCAGAATAAACC
CTGGTAGTCAATGGGGTGACAGTGTCGCAGCCAGATTGCCCTCACATCCAACTCTTAGTGATCTTCTCTTAACATTTCTTGCAAGGCAG

FIG.3A-1

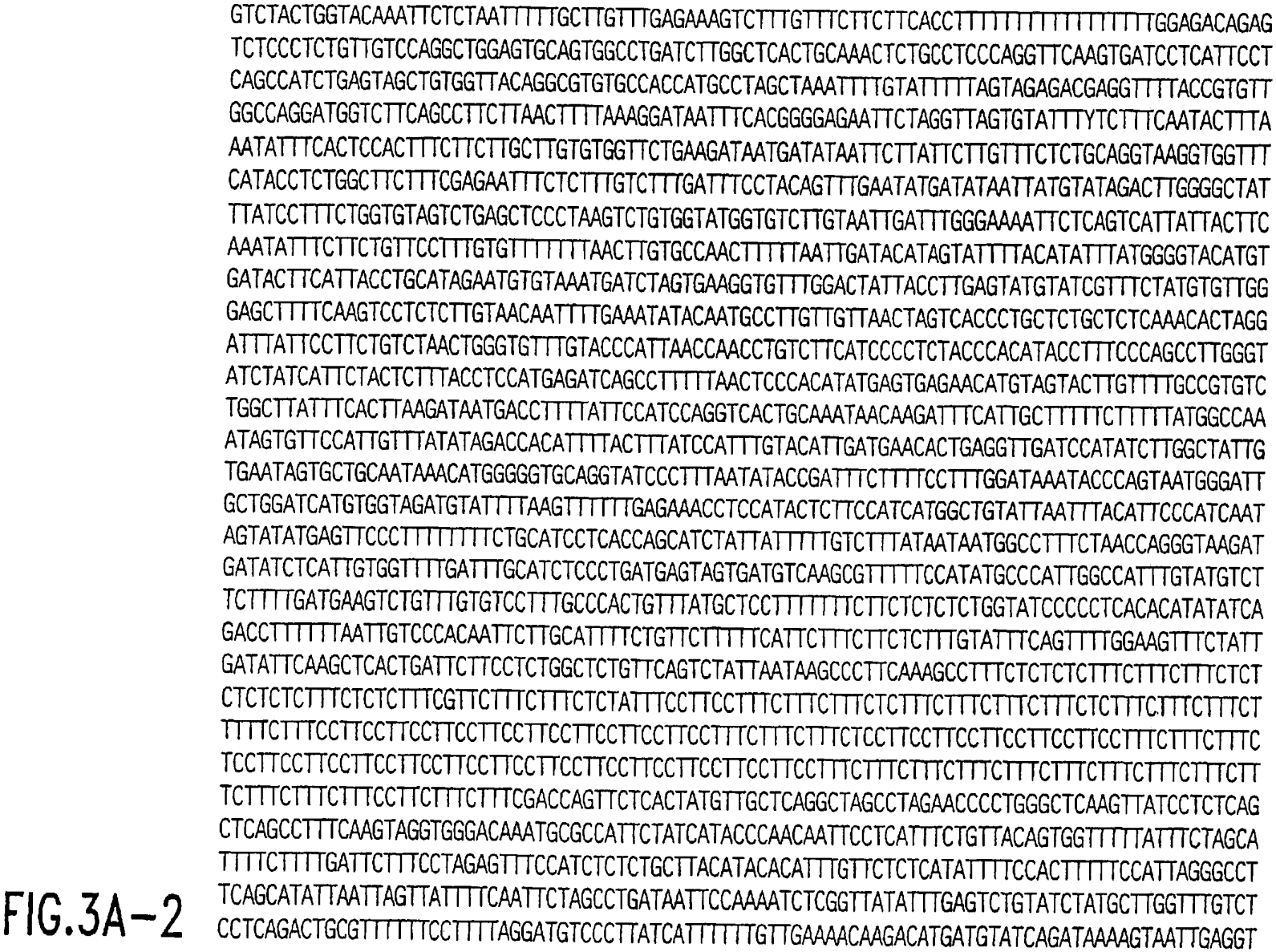

GTCTACTGGTACAAATTCTCTAATTTTTGCTTGTTTGAGAAAGTCTTTGTTTCTTCTTCACCTTTTTTTTTTTTTTTTTTGGAGACAGAG
TCTCCCTCTGTTGTCCAGGCTGGAGTGCAGTGGCCTGATCTTGGCTCACTGCAAACTCTGCCTCCCAGGTTCAAGTGATCCTCATTCCT
CAGCCATCTGAGTAGCTGTGGTTACAGGCGTGTGCCACCATGCCTAGCTAAATTTTGTATTTTTAGTAGAGACGAGGTTTTACCGTGTT
GGCCAGGATGGTCTTCAGCCTTCTTAACTTTTAAAGGATAATTTCACGGGGAGAATTCTAGGTTAGTGTATTTYTCTTTCAATACTTTA
AATATTTCACTCCACTTTCTTCTTGCTTGTGTGGTTCTGAAGATAATGATATAATTCTTATTCTTGTTTCTCTGCAGGTAAGGTGGTTT
CATACCTCTGGCTTCTTTCGAGAATTTCTCTTTGTCTTTGATTTCCTACAGTTTGAATATGATATAATTATGTATAGACTTGGGGCTAT
TTATCCTTTCTGGTGTAGTCTGAGCTCCCTAAGTCTGTGGTATGGTGTCTTGTAATTGATTTGGGAAAATTCTCAGTCATTATTACTTC
AAATATTTCTTCTGTTCCTTTGTGTTTTTTTAACTTGTGCCAACTTTTTAATTGATACATAGTATTTTACATATTTATGGGGTACATGT
GATACTTCATTACCTGCATAGAATGTGTAAATGATCTAGTGAAGGTGTTTGGACTATTACCTTGAGTATGTATCGTTTCTATGTGTTGG
GAGCTTTTCAAGTCCTCTCTTGTAACAATTTTGAAATATACAATGCCTTGTTGTTAACTAGTCACCCTGCTCTGCTCTCAAACACTAGG
ATTTATTCCTTCTGTCTAACTGGGTGTTTGTACCCATTAACCAACCTGTCTTCATCCCCTCTACCCACATACCTTTCCCAGCCTTGGGT
ATCTATCATTCTACTCTTTACCTCCATGAGATCAGCCTTTTTAACTCCCACATATGAGTGAGAACATGTAGTACTTGTTTTGCCGTGTC
TGGCTTATTTCACTTAAGATAATGACCTTTTATTCCATCCAGGTCACTGCAAATAACAAGATTTCATTGCTTTTTCTTTTTATGGCCAA
ATAGTGTTCCATTGTTTATATAGACCACATTTTACTTTATCCATTTGTACATTGATGAACACTGAGGTTGATCCATATCTTGGCTATTG
TGAATAGTGCTGCAATAAACATGGGGGTGCAGGTATCCCTTTAATATACCGATTTCTTTTCCTTTGGATAAATACCCAGTAATGGGATT
GCTGGATCATGTGGTAGATGTATTTTAAGTTTTTTGAGAAACCTCCATACTCTTCCATCATGGCTGTATTAATTTACATTCCCATCAAT
AGTATATGAGTTCCCTTTTTTTTCTGCATCCTCACCAGCATCTATTATTTTTGTCTTTATAATAATGGCCTTTCTAACCAGGGTAAGAT
GATATCTCATTGTGGTTTTGATTTGCATCTCCCTGATGAGTAGTGATGTCAAGCGTTTTTCCATATGCCCATTGGCCATTTGTATGTCT
TCTTTTGATGAAGTCTGTTTGTGTCCTTTGCCCACTGTTTATGCTCCTTTTTTTCTTCTCTCTCTGGTATCCCCCTCACACATATATCA
GACCTTTTTTAATTGTCCCACAATTCTTGCATTTTCTGTTCTTTTTCATTCTTTCTTCTCTTTGTATTTCAGTTTTGGAAGTTTCTATT
GATATTCAAGCTCACTGATTCTTCCTCTGGCTCTGTTCAGTCTATTAATAAGCCCTTCAAAGCCTTTCTCTCTCTTTCTTTCTTTCTCT
CTCTCTCTTTCTCTCTTTCGTTCTTTCTTTCTCTATTTCCTTCCTTTCTTTCTTTCTCTTTCTTTCTTTCTTTCTCTTTCTTTCTTTCT
TTTTCTTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTTCTTTCTTTCTCCTTCCTTCCTTCCTTCCTTCCTTTCTTTCTTTC
TCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTT
TCTTTCTTTCTTTCCTTCTTTCTTTCGACCAGTTCTCACTATGTTGCTCAGGCTAGCCTAGAACCCCTGGGCTCAAGTTATCCTCTCAG
CTCAGCCTTTCAAGTAGGTGGGACAAATGCGCCATTCTATCATACCCAACAATTCCTCATTTCTGTTACAGTGGTTTTTATTTCTAGCA
TTTTCTTTTGATTCTTTCCTAGAGTTTCCATCTCTCTGCTTACATACACATTTGTTCTCTCATATTTTCCACTTTTTCCATTAGGGCCT
TCAGCATATTAATTAGTTATTTTCAATTCTAGCCTGATAATTCCAAAATCTCGGTTATATTTGAGTCTGTATCTATGCTTGGTTTGTCT
CCTCAGACTGCGTTTTTTCCTTTTAGGATGTCCCTTATCATTTTTTGTTGAAAACAAGACATGATGTATCAGATAAAAGTAATTGAGGT

FIG.3A-2

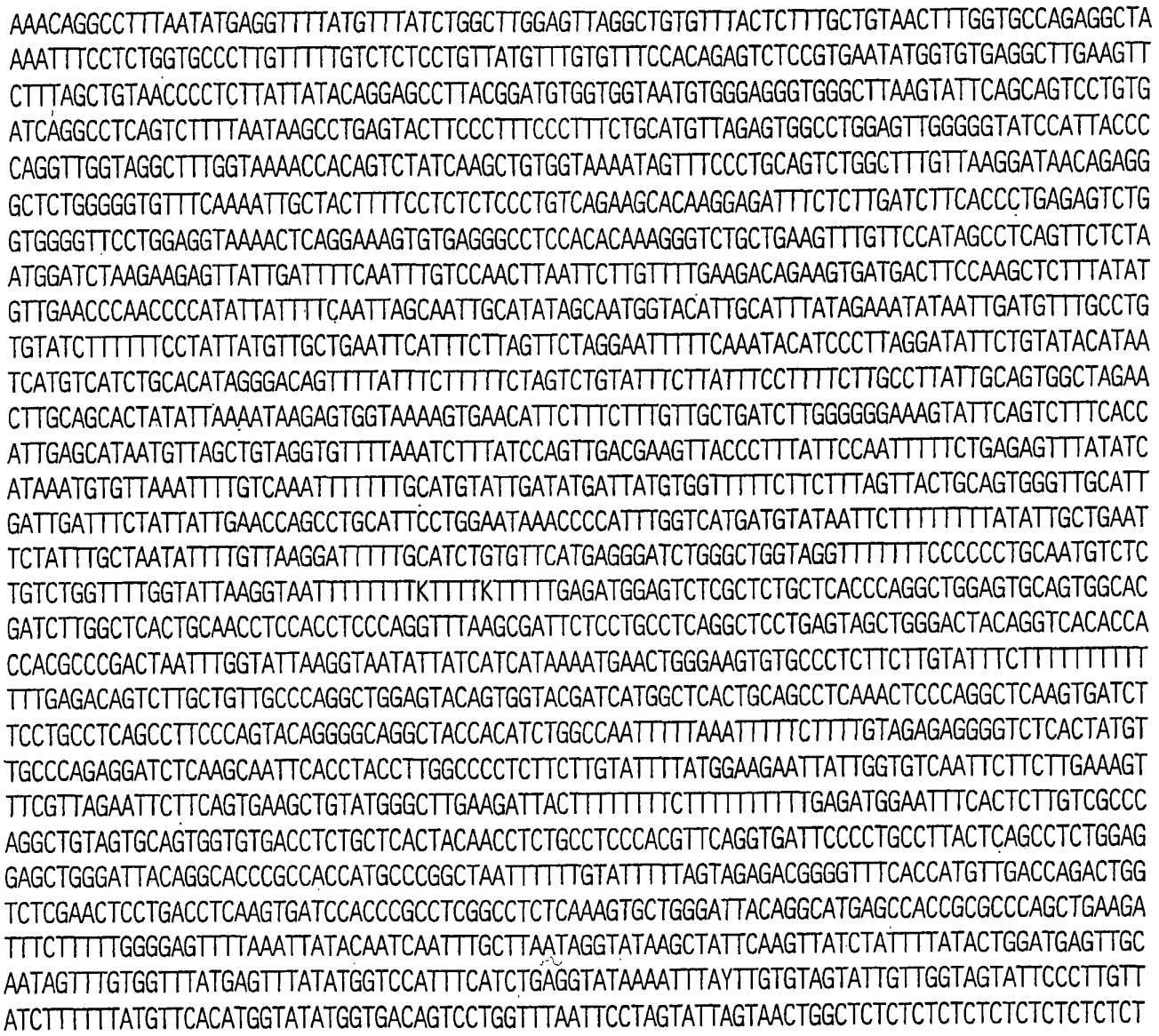


FIG.3A-3

CTCTCTCTCTGGTCAGTCTTTCCAGAGGTTTGTCAATTTTGTTGACTTTTTTCCCCCAAAGAATCAGCTCTTTGTTTCATGGATTTTCT
GCTTTTCTGTTTTCAACTTCATTGATTTCTGCTGTTTATTATTTCTCTCCTTCTGTTGGTTGTGAGTTTGTTTTGCTTTTCTTTTTCTA
CATATTCGATGTGAAATCTTACATTATTCACTCGGGACTTTTCTTCTTTTTTGATGTATGCATTTAGTATTCTAAATTTACTTCTKAGT
ACTGCATACTGCTTGAACTATGTCTGACAAATATTAATATATTGTTTTTAAATCTTTATTCAGTTCAGTGTATTTTTAAAATTTCCTTC
TCTGCCTCTTCTTTGATTTGTTATTTAGAATTGTGTTGTTATTTTCCGAGTATTTACATTTTCCTCTTATCTTTCTGCATTGATTCCAT
CGTAGTCAGAGTGCATGCTCTGTACAGTTTCAGTTCTTTCAAATTTATTGAGCTTTGTTTAATGGATCTGGATACAGTTTATCTTGGCA
TATATATATATATACACACACATATGTATGTGGGCGCTTGAAAAGAAAGCGTATCTGCTGTTTGGTGGAATGTTTGGAGTGTTCTATAA
GCGGTGATTAGATACTGTTGGTTGATGATGTCATTGAGGGTCCGATAACCCTACTGATTTAAATTTATTTAGTCTGTCAATTATTCAGA
GAGAGAGGTGTTGAACTCTGCAATGTGAATTGTGGATTTGTCAATTTCTCCTTTCAGTTCTATTAGTTTTTTCTTCACATATTTTACAA
CTCTGTTGTTTGGTGCATACACATTTATGCACCAAATTTAGGATTGCTATAACTTCTTGGTGGATTGACCCTTTTACATTATATAATGT
CTTTTTCTGTCCCTGGTAATTGTGGTTGCTCTGAAGTCTATGTTATCTCAATATAAATAGACAACTCTGCTTTCTTTTGATTAATGTTT
ACATGATACATCTTTTTCTATTCTTTTACTTTCAACTTACTTATATTATTATGTTTGAAGTGAGCTTCTTGTAGACAGCATGTAGTAGG
TCATATATGTACATAGATATATATATTTTTTTGAGATGGTGTACTCTGTCACCCAGGCTGGAGTACAGTAGTGCTCACTGCAACCTCTG
CCTCCTGGGTCAAGTGATCTCGTGCCKCAGCCKCCCCAGTAGCTGGGATTACAGGCACGCACCACCATGCCCAGCTAATTTTTGTATTT
TTAGTAGAGACGGGTTTAACCATGATGGACAGGCTGGTCTCGAACTCCCGACCTCCAGCGATTAGCCCACCTTGGCCTCCCAAAGTGCT
GGCATTACAGGTGTGAGCCACCGTGCCTGGTTTAATATTTTTAATCCACTCAGTCTTTGTCTTCTACTGGTGTACATAGACATTCGCAT
GTAATGTAAATGTTGATATGTAAGAGCTTGAATCTGTTATGTTTTTGCTTTCTCTATGTTTTCTCAATTTTTAATTTCTCTGTTTTCTT
TTTTTCTGCTTCATATTGGCTAATGAACACTTTGAATCATTCCATTTTGATTTACCTATAGTGTTTTTTAGTGTGTCTCTTTGCATAGC
TTTTTTAGGGGTTACTTTAAGTATTTCATTATATGTACATAACTTATCACAGTATATTGGTATCGTTATTTTACCAGTTCAAGTAAAGT
ATGGAAATGTTTCCTCTCTACATTCCTTTACCTCATTTATAATATAATTGTCTTAGGTATTTCTTGTACATACATTTTAAACCGGATGA
GTGTTATTTTTGATTTAGCTATCAAATAATTCCAAAAACTCAAGAAAAAAAGGAAAGCTTACTATATTGACCCATATTTTCATTCACCA
TGTTGTTTCTTCCCTCTTTATGCCCCATAGTTCCTTCTTCTATTGTTTTCGTTTAGAGAACTTCCTAGCCATTCTATTGGGGTAGATCT
CCTAGTGACAAATTCTCTTAGCTTTCTTTTCTCTGTGAATGTCTTTATTTCCCTCTTTGTTCCTGGAGGACATTCTCACTGGATATAGG
ATTCTTGGCTATTGGGTCTTTTCTTTTGGCACTTTTGTAAGTGTGCAGCCTGCTGTCAAAATAAAAATTAAAATAAAATAAAAATGAAT
GTTTTCCTTTGCTACGTTCATGAAAGTATAATTCACTGAATGAGGAGGGACACCCATCTCTATAATCTGGAGGCCCATGCTCACCTCTG
AATAGTACATTTGCAGAGAAATTGGGGAAATCAAAGTCTGTTGAGACCAGCAAGATAAATAAGGCAAAAGGATACAAAACCATATCCAA
AGAGAAATGGTTTAAAGGAACTAAGGCTGTTTCTCCTAAAAAGAAAATAGTTGGAGACATGTGACCTCCAAAGAAACAGGACTTTTTCT
ATGGGGCTCCAAGGGGTTTCTATGAGAGAATGATAAAGGAGAGATTTCAGCTTAGTCTCAGGAAGACTTTTCAACAACCAAACCTGCCC
AAAGATGGACTGCCCTGCCTAAGGATTGTGTTCTGACATTAAGGGTATGGAGGTATGGGTTAGATGAATATTTTACCAAAATGCCATAG

FIG.3A-4

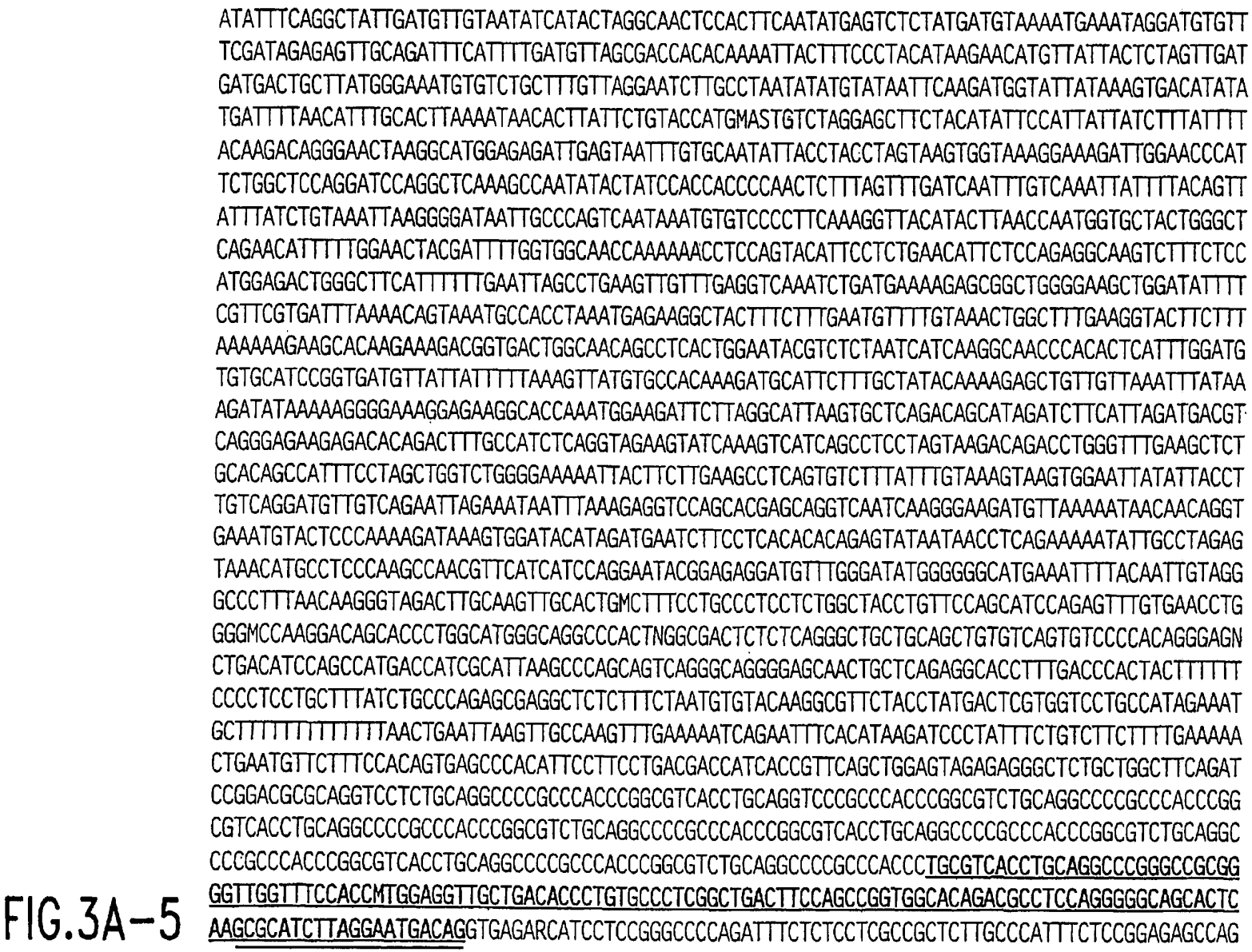

ATATTTCAGGCTATTGATGTTGTAATATCATACTAGGCAACTCCACTTCAATATGAGTCTCTATGATGTAAAATGAAATAGGATGTGTT
TCGATAGAGAGTTGCAGATTTCATTTTGATGTTAGCGACCACACAAAATTACTTTCCCTACATAAGAACATGTTATTACTCTAGTTGAT
GATGACTGCTTATGGGAAATGTGTCTGCTTTGTTAGGAATCTTGCCTAATATATGTATAATTCAAGATGGTATTATAAAGTGACATATA
TGATTTTAACATTTGCACTTAAAATAACACTTATTCTGTACCATGMASTGTCTAGGAGCTTCTACATATTCCATTATTATCTTTATTTT
ACAAGACAGGGAACTAAGGCATGGAGAGATTGAGTAATTTGTGCAATATTACCTACCTAGTAAGTGGTAAAGGAAAGATTGGAACCCAT
TCTGGCTCCAGGATCCAGGCTCAAAGCCAATATACTATCCACCACCCCAACTCTTTAGTTTGATCAATTTGTCAAATTATTTTACAGTT
ATTTATCTGTAAATTAAGGGGATAATTGCCCAGTCAATAAATGTGTCCCCTTCAAAGGTTACATACTTAACCAATGGTGCTACTGGGCT
CAGAACATTTTTGGAACTACGATTTTGGTGGCAACCAAAAAACCTCCAGTACATTCCTCTGAACATTCTCCAGAGGCAAGTCTTTCTCC
ATGGAGACTGGGCTTCATTTTTTGAATTAGCCTGAAGTTGTTTGAGGTCAAATCTGATGAAAAGAGCGGCTGGGGAAGCTGGATATTTT
CGTTCGTGATTTAAAACAGTAAATGCCACCTAAATGAGAAGGCTACTTTCTTTGAATGTTTTGTAAACTGGCTTTGAAGGTACTTCTTT
AAAAAAGAAGCACAAGAAAGACGGTGACTGGCAACAGCCTCACTGGAATACGTCTCTAATCATCAAGGCAACCCACACTCATTTGGATG
TGTGCATCCGGTGATGTTATTATTTTTAAAGTTATGTGCCACAAAGATGCATTCTTTGCTATACAAAAGAGCTGTTGTTAAATTTATAA
AGATATAAAAAGGGGAAAGGAGAAGGCACCAAATGGAAGATTCTTAGGCATTAAGTGCTCAGACAGCATAGATCTTCATTAGATGACGT
CAGGGAGAAGAGACACAGACTTTGCCATCTCAGGTAGAAGTATCAAAGTCATCAGCCTCCTAGTAAGACAGACCTGGGTTTGAAGCTCT
GCACAGCCATTTCCTAGCTGGTCTGGGGAAAAATTACTTCTTGAAGCCTCAGTGTCTTTATTTGTAAAGTAAGTGGAATTATATTACCT
TGTCAGGATGTTGTCAGAATTAGAAATAATTTAAAGAGGTCCAGCACGAGCAGGTCAATCAAGGGAAGATGTTAAAAATAACAACAGGT
GAAATGTACTCCCAAAAGATAAAGTGGATACATAGATGAATCTTCCTCACACACAGAGTATAATAACCTCAGAAAAATATTGCCTAGAG
TAAACATGCCTCCCAAGCCAACGTTCATCATCCAGGAATACGGAGAGGATGTTTGGGATATGGGGGGCATGAAATTTTACAATTGTAGG
GCCCTTTAACAAGGGTAGACTTGCAAGTTGCACTGMCTTTCCTGCCCTCCTCTGGCTACCTGTTCCAGCATCCAGAGTTTGTGAACCTG
GGGMCCAAGGACAGCACCCTGGCATGGGCAGGCCCACTNGGCGACTCTCTCAGGGCTGCTGCAGCTGTGTCAGTGTCCCCACAGGGAGN
CTGACATCCAGCCATGACCATCGCATTAAGCCCAGCAGTCAGGGCAGGGGAGCAACTGCTCAGAGGCACCTTTGACCCACTACTTTTTT
CCCCTCCTGCTTTATCTGCCCAGAGCGAGGCTCTCTTTCTAATGTGTACAAGGCGTTCTACCTATGACTCGTGGTCCTGCCATAGAAAT
GCTTTTTTTTTTTTTTAACTGAATTAAGTTGCCAAGTTTGAAAAATCAGAATTTCACATAAGATCCCTATTTCTGTCTTCTTTTGAAAAA
CTGAATGTTCTTTCCACAGTGAGCCCACATTCCTTCCTGACGACCATCACCGTTCAGCTGGAGTAGAGAGGGCTCTGCTGGCTTCAGAT
CCGGACGCGCAGGTCCTCTGCAGGCCCCGCCCACCCGGCGTCACCTGCAGGTCCCGCCCACCCGGCGTCTGCAGGCCCCGCCCACCCGG
CGTCACCTGCAGGCCCCGCCCACCCGGCGTCTGCAGGCCCCGCCCACCCGGCGTCACCTGCAGGCCCCGCCCACCCGGCGTCTGCAGGC
CCCGCCCACCCGGCGTCACCTGCAGGCCCCGCCCACCCGGCGTCTGCAGGCCCCGCCCACCCTGCGTCACCTGCAGGCCCGGGCCGCGG
GGTTGGTTTCCACCMTGGAGGTTGCTGACACCCTGTGCCCTCGGCTGACTTCCAGCCGGTGGCACAGACGCCTCCAGGGGGCAGCACTC
AAGCGCATCTTAGGAATGACAGGTGAGARCATCCTCCGGGCCCCAGATTTCTCTCCTCGCCGCTCTTGCCCATTTCTCCGGAGAGCCAG

FIG.3A-5

AGAAAGCCGCTCCCAAGTCCAAGGCCGAGCTCCGCAGACGCCCGGCCCCTCCGGCGCGGACAGAACAAAGCCATTGTTCTTGCCGGGGA
AGGTAGAAATACTGTGGGCTGCTTCAGAGGCTGCCGAGCAAAACTCAGGCAATCTCCTGGGCTGTTCCAATACGTTTATTCTCTTTTTC
AAAACAGGAGGAGGAGGTAGAGGCGGGGAGACACACCATCCCTGCAAAACTACTGGCAAAAACTAAGCGGAGCCGGGTGTGGTGGCTCA
CGCCTGTAATCTCAACACTTTGGGAGGCCGAGGGGGGCCGATCACTTGAGGTCAGGAGTTGGAGGCCAGCCTGGCCGGCATGGTGAAAC
ACAAAAATTAGTCGATTGTGGTGGTGCATGCTTGTAATCCCATCTACTTGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCG
GAGGTTGCAGTGAGCCGAGATTGCGCCACTGCACTCCAGCCTGGACAACACAAGTGAGATTCTGTCTCAAAATAAATAAATAAATAAAC
CCAAGCAGAAAAAGAATCACTCTGAAAACGATCACATCTAACTATCAATGCTCATACAGTTTATGGAATTATCAGCCCAACTTGATAAA
ATCAGTATTTGAGGAAACTGTGGATAAGCCCCCTGATTTCAATCCCCATTGTGCCAGGTCCTGGTTAACTGAGGTTAACGAAGTAAAGA
GCTGCAGACACTATTAACTGCTACCTTAAACCGATTACTCTAGCTTAGCCTACTTTCCACGTACAGATTTTACCAGTGGACAACATGAT
GCTTTATCTTGTTTTTCTCTCCCTGGGACTTTTCTCCAGACATTGAAAACAGAAATACTAATAAGGCCACTTTTACCTGCCTGATGCAA
GAACAGAATTTTCAAACTCAACATTAATGCAACTCCTCAGTCCCTGACAATGGCGGGTGGAAAAGTTTCTAAAAATATGCAGCAGCACA
ATTATCGGGAAGAGATGAGATACTGTTACCTAATAAAAATGCCATAAATAGAGAATGATGAACTACCATGGGAAATGAATGCATAGAAG
AGGACATGCTGGAATGTGGGACAGTAAAAATCACTTAAACTTTGCGTGACCTTGAAGAAAGTCACGATGATCTGTTTTTCCAGGTCCCT
CAAACAGTGAGATGTGGCTGTTTCCCAAGTCTTCCTCTCCAGTGTAAAGGGTCTGAATTTAGACGCTTTGTGAGTCTTCCTTCTTTCGA
CAGCCTGGAGTCTCTCTTGAGTCTCAAGGCTGCCTGAGTTCCTCTCTAACATCCTCTAGGCAGTATCAGCTAATGAGACAATGAATTCC
ATGGAGGCAGCAGTGGGAACAGAAGTACCTCTCTTGGATAATTTACAACACTGGTGAGCAGAGGGTCAGATCACCCTGGGGTTTGTGTC
ACAACCAAAAAAGTGGCTGTGGCACTGAGTTCTTGGATGGTTTTCTACAGCTGGTCCAGATTTTCCATGGGCTCACCTTTAAATTAAAA
GAATTTCTGCACTTTGAAGAATTTGAAAACAAAGCCATGTGTGAGAATATGAGATCCACTCATATGCCCTTGCAAGAAATAGGTTGCAT
TCCTTTTTCCGGACTTAAAAAAAAAGCACCCCCTCTTTCTTTTTTTCAGAAGGCATATATGTAAATGATTCCAAATTAATCTTTAGCAT
GTGCCTATGTTGTTCTGATTTACTAAACTTTAAAAATATGTCCATTGTTGTCTGTTAACAGCTTTTGGCAACTTTTTCAGAGATTGAAA
TATGTGAGCAAATTAGAGAAATGAGTACAATTATTAGCTAGTACCATTCAACAAGCGCTAAAGATACAAATACCTCTACAATACATAAA
AGGAATGATTATAGTAGATTTTATAATGCCATATAAGGTTTCTTATTTAACTTCATTCTTAATTCTCAAAATAAAATGAAATTACATAG
AAGCAAAGTAATATAGTTACCAGAATAGTATTTTTACATGTCTTTAAGTGTATGTTGTTGTTGTTGTTTTTAAGGTAATTATGTGATGT
TGTGGAAAGAACAGAGACCTGGGTTAGATAAAATTCCGGTTGTCTACCAGATTGTGATAGTGAGCAAATTACTTAACCTCTATGATCCT
TATCTTATTTATCTATGAAACAGGATTGGTAATACTCATATCATAAGGTTGAAAGGATTAAATGAGGCACTATGGAAAATTTCTAACAT
GGTGGTGCCTGGGACAGTAGAAGATGCTTAATAAAGATAGCTTTCATTATTATTATTAGCTTTTTCAGGTGATGGTGATTGTAAATGTT
TAGGTAATTTTTTTAAACTTTAGAAATAATTGATTTTCAAATGATTAAGACTGCTTATTTTAATCATTTATTTTTATCACCAGATTTATT
TTTATTACCCAAAATGTCAACGACTGTCATAAAGATAAAAATTAATAATAATTGGCCAGGTGCGGTGGTTCACGCCTGTAATCCCAGCA
CTTTGGGAGCTGAGGTGGGTAGATCACAAGGTCAGGAGATTGAGACCATCCTGGCTAACGCGGTGAAACCCCATCTCTACTAAAAATAC

FIG.3A-6

AAAAAATTAGCTGGGTGTGTTGGCGGGCGCCTGTAGTCCCAGCTACTCATAGTCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATGGTG
TGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCGCGCCACTGCACTCCAGTCTGGGCTACAGAGCCAGACTTCATCTCAAAAAAA
AAAAAAAATTAATAATAATTTAAACCCGAAGTATGAACTGAATTATTTCCCTTAGTAGCACATCACATAGGCTGATGATAGTTTTGGTG
ACTGGTTTATCTATTCTTCCTAAAAGCAAACTGTTGTTAGATGGATGATCACTTGCATGTTGTGACTGAACTCAGCAGTTGGGTTTTAT
TTTTTATTTTTTATTTGCTTCAGTAGCATTADCCTTTCCTACCAAGATTCGAACAATCCATTTGCCTTTTTTTCCCTAAAATCTCTCAT
ACATTGTAAATACTACATATTGGCTAAATATTTCCTGGACAGACATGAAGGACACATAAATCAGTCTCTGTATGATGTTTCTCACTGTA
ATGGAGTTTATCTGGCTCAAGACCAGGACATTTATTGCATATCAGGTTTCTACAGTTCAGGCAAAAGTTTGAGGATAAGGACTTACTGC
AAAAAGTCTTCTATTGTTCTCAACCATTTTCTCGCTTAGCACATGCAGAGATTTGAAATGGTCCGTGGTACAGTAGTTGTGTCTGTATA
TTTCTCTTGTAGAATATTAGAACAAGGGATTTGCAGTTTACAGAGAAGAAGGCTTGGCGAGGTGTTTGGAAATACACTCAGAAACCTGA
GGAAATTTGTGGAAAGAGAGGCTTATTATTTCTAGAAATATGCTAGAGTWCGTTTTGATTGTGCACCTGAGGAATTAATAGATTAAGTA
GTTTTATAAGGACTGGGGTTAATAGAATACTGGCAGTGAAGTTTGTCTTAGGACTTCTTAATTGGATAATCAGTGAAGTCACCAGATCC
CAGTTAGAGACAGTTCCAAGTTTTACAAAACGCAAGATAACTGTCCAAGAGCTGTAATGGCTTAATCATCTTTGAATAATACCTCTCAC
TGAAGCTATATCATAAGAAATAAAAATCTACATTTTAAAAAATTGGCTGTAATCATAGGGTGACTAACTGTCCCTGTTTACCCAGGACT
CAGGGTTTCCCAGGCTGAGGGACAATGGGTACTAAAACCAGGACAGTCCCAGGCAAACTGGGACGGTTGATCACCCTACCCAATGGCCT
CATCTGTCTCATTAAAATATCTGGATTACTTCGTGCCTCAAAAATATCCTCGGCTTACCTGACTCTAGACAGTCAAGAAGCTTTTATTA
ATTGTCTAATGTATGCCACTTTCTGGAGGTGATATTGTTCAACTGATAGATGAGCATCACTGATTGAAATATTTTGTGGTTTTCATGCT
TTGTATCTTGTGCTGATAGCCCCACATGGATATTTCTGTTTCCAAGTTTGTGTCACTTCTGGAGATATTAGCCTGAACTCAGCAAAATA
GGATGATCAAAATGAACCTTTCCAGTGAATTCTGTCCTTCTTGTGCTGTTGTCATCTGACTTAGATATACTGGCCGGGCGCGGTGGCTC
ACACCTGTAATCCCAGTACTTTGGGAGGCTGAGGTGGTTGGATCCCTTGGGATCAGGAGTTTGAGACCAGCCTGGCCAATATGGTGAAT
GAAGCCCTGTCTCTACTAAAAATACAAAAATTAGTTGTGCGTGGTGAAGTGTGCCTGTAATCCCAGGTACTCAGGAGGTTGAGGCAGGA
GAACTGCTTGAACCAGGGAGTCGGAGGTTGCAGTGAGCCCAGATCACACCACTGCACTCCAGCCTGGCAACAGAGTGAGACTCCATCTC
AAAAAAAAAAAAAATTAGCTGGATGTGGTGGCACATGCCTGTAATCCCAGCTACCTGGAAGGCTGAGGCAGGAGAATCGCTTGAACCCA
GGAGACGGAGGTTGCAGTGGGACGAGATCGTGCCACTGCACTCCAGCCTGGGTGTCACAGCGAGACTCCATCTCAAAAAATAAAAATCA
ATAAAAAATAAATAAATACATAAATAAATGAACACATAAATTAGATATACCAAGAAAAGTATAAAAAAGTCTTGTGTGAACATAAATGA
AAATTGGCCAAAATAGGTAACAGACAGGGTCAGGCGTGGTGGCTCATGCCTGGAATCCCAGTACTTTGGGAGGCTGAGGTGGGAGGACC
ACTTGAGGCCAGGAGCTCAAGACCAGCTTGGGCAACAAAGCGAGACCTCATCTCTATGAAAGAAAAAAAAATTTAAAAGACGTAATGAA
CAACTTGCTTGCCTTCCTGCCTGCCTTCCCTAAAATACTAAGTTAAATGCAATACATGCCCTGACATTGTAGTTTGCTTTCACAAAGAT
TTACTGAATACTTACTCTAGGCTAAACCTTGTGCTACATGTTGGGGCTACAGGGATGAAAGARAATTGGTCTTGCCCTCCAGGAACCTT
TCATTTAGTACAGAGATTTAGTGTGTGCTGGTTGGTCTCTGTTCTCCCCCTCTCCTCCAGATCTATTCTCTATTTCTTCCCCTCTCCCT

FIG.3A-7

GCCTCCAGGAAGGGGGGCTGGATCACTGTGGCTCATTGCTCTGTGGCTTCTGATTGAGTTCAGCCAATGGGAGGCATMATTTTGGCGTG
GCAGCTCTGGCTGTTCCTCTGCAATTGCAGTTCCCTCCTCCAAGGCTCTGGCTCTCACTGGGTTCCTGTATCCAATAACAGACTCCCTT
AACTGCCCACTTCTGAAAACAGTTTCTGCATAAAGCTATTTTCATAATTTCCTCTGATGTGCCTTCTGTTTCCTGTGTAGACCCTGATT
CAATAGGAAAATAAATTATTGAAATAGAGGAAGAGACAGGTAATAATAGAGGTATACACAAGTAGAATGGGGCAATAAATGGCGCATTT
TCGCACCATCAAGAGTGCCCATGTAACAGAGATAAGTAAATGCATCTTGAGCTGAACACTGAAGGATAAGAAACAAAGGGGAGAAAGAC
CTAGAAGGGGCAATATACAGCAAGGAGGGAAAATAAACTACTGTGCATTCATGCCAGTGTTAGCATTTAGGACATCTGGAAGCTAGAGG
TGGAGTGGAAAAGGAGAGAGTGATAGGAGCTGGGGTCAGAGAGTTTCAGGGTGGGGAAGGTCTTGCAGGACCTTGTAGGTAATTGTAAA
GCATTTGGATTTTATTCTGAGGGTCACTGGGGTGTCATTAGAGACTTTTGAGCAAAGAGGTACATGCTCTGACTGAACTTTATTCTGTG
AACAATCAGAATCAACTAGATGGATTTAAGTATGGGTATACCATGAAAGAAAATTACTTAAGATCCTTGCTACTCAAAGTATGAGCCAG
GACCAGCTACACTGGCATMAGCTGGGAACTTGTTAGAAATGCAGAATCCCAAGTCCCCGAGACAAACTGAATCAGAACCTGCACTTTAA
CAAGATCCCAGGTGGCCCATTTGTATGGTAGAGTTTAAGAAGCATTGGTTTAAAAGATCCCTCTTGATAGGAGCATGGAAGATACATTT
GAGACAGAATAGACAAGTCAGAGACAGGTGGGAAGGGCCTAAAACAGGGCAGAAGTAGGGAGGTAAATGAGGAGACAAATACAAAGGAA
GAAAATGCACAGCACAGTGTAGACAATTCCTAAATACTTAAAAAAATTTTTTTTGAAATAATGATAGATTCACAGGAGGTTGCAAAGAA
ATGCGTAGGGAAGAACAATGCACCCTTTACCCAGCCTCCTCCATCATTAACATCTTATGCAACTATATTATAATATCGAAAACAATCAA
GTGACATTGCTACAACCCATAGAGCTTATTCAGATTTCACCAGTTATTAGATGCACTCGTGTGTGTGTATGCATATAGCTCTGTGTAAT
TTTATCATATGTGAAGCTTTGCTACCACAATCAAGATATTCAAGCCATTAGCAGAAGATTTTCTGGTGTTACCTCCTTATAGCCACACG
CATTCCTCCATCATTAACCCCTGGGAACAACTAATCTGTTCATCTCTATAATTATTCTATTTCACGAACATTTTGTAGATGGGTACATG
CAGTGTGTATCTTTTGGGATTGGTAACAGAGCAAGACAGGATCTCACTCTGTCACCCAGGCTGGAGTGCAGTGTCGTGATCTTGGCTCA
TTGCAGCCTCCACCTCCTGGGCTCAGGTGATCCTTCCACCCCAGCCTCCTGAGTAGCTGGGACTACAGACACACGCCACCTCACCTGGC
TAATTTTTTGTATTTTTATAATGATGGGGTTTCACCATTTTGCCTAGGCTAGTCTAGAACTCCTGGGCTCAAGTGATCCAACCGCCTTG
GCCTCCCAAAATGCTGGGAGTACAGGCATGAGCCACCACCTCCACCAGCTTTTTCATTCATACTTTCTTTGAAGTTCATCCAAGTTGTG
TGTATCAATACTTCACTCCTTCCAGTTGCTGAGTAGTATTCCATGGCTTGGAGGTGCTAGAGTTTATTCATCACATTCAACCCATTGAA
GGMCATTTGGGTGGCTTCCAAGTTTCCAGTTTTGGGCTATTATGAACAAAGTTACTATGAACATTCATATACAATGGATACTTTTTGTA
TGAATGAATGGAATAGAATGGATAGGATTTAGTGATCAGCTATGTGGGATGAAGAGTGGCATAAGTAGTAAAAAGTAACCCTCAATGCA
ATGTGCAGCCAGCAAGTACCACAAAAAGAGTTTATTTTGTTTCATACATATATTTCTATATATACATACACACACTTTATTAATAACCA
AATAGTATCCTTTTCAAATGAAAACAGTAATTTAACATAAACTATGAACTTAAAATCTAAAGTAAAACTTGACAACAGTGATGCAGAAT
TTTTTGCTCCTTAGCTCAGTTAGGTCTGTGTTCTTATCTTATGACCAGGAAGAACTAGGTACCCTGACATCAAAGAATGAGTGGCATAG
AATTTATTAAGCAAAAAGGAAAGCTCTCAGGAAAGAGTGGGGTCCTGAAAGCAGGTTGCTGGTTGCCCCTTCGTAGTTGAATACAAGGG
CTTCTATATAAAACCTGATGGGGCCGAGTTCCCTGTTCGTATAAGGCATGAATTCCTGGTGGCTCCACCGCCCTCCCCCAGTGCGTATG

FIG.3A-8

TGGGACCTTCGTCCACTAGGGACATGTTTAGACAAGCTCCCTGTGCACGTTCCCTTATCTGCACAAAACATGGGTTGGAGGTTCTCCGG
GGACCCTTCCTTTACTTTCTGCCTAAAGCAAGCTGGCTAACTCCTTTCAACAATACTAAAGACATACAGACAATGGTTCTCAGTACAAT
CATTTTAAATATTTAAGTAAACTTAAAATGGTGTTTGTTTTGATTTGACATTTTAAAAGATATCGCTGTTCTAAAAATTCTGTGTTTTT
AGTTGTTTGGGCTCCTATTCTACAATGTGCTATTACTATTAAGCATTCTTGTATCATGGCATTCCTCAAATAGTTTTTAAATTACTTTT
AATTTGAAGAAGGAACATTCTGTACAGTCACGGAAAGTGTCAAAAATGAAAATGAGGCAGGGTGTGGTGGCTCACGCCTGTAATCTCCG
CACTTTGGGAGGCCTAGGTGGGTGGATTGCTTGAGCCTAAGAATTTGAGACCAGCCTGGGCAATATGGTATAACCCTGTGTGTACAAAA
AATACAAAAATTAGCCAGGTGTGGTGGCCCAAGCCTGTAGTCCCAGCTACTTGGGAAGTTAGGGTGGGAAATCCTAGGTGACAGAATGA
GACCTTGTCTCAAAAAAAAAAAGAAAAAAGAAAATGATAAAGGATACATATCAGGAAAACATGCATGGTATTTTGTATCATCTACTTTA
GAGTAATTCCAGTATAGTGGTTTTTTTGTTGTTGTTTGTTTTATTTTTGAGAAAGGGTCTTGCGCTGTCACCCAGGCTGGAGTGCAGTG
GTACGATCTTGGCTCACTGCAACCTCCGCCTACCAGGTTCAAGCCATCCTCCCAACTCAGCCTCCAGAGTAGCTGGGACTACAGGTGTG
CGCCACCATGTCCAGATAATTTTGTATTTTTTGTAGAGATGGGATTTTGCCATGTTGCCTGAATGCCTGGCCTCAAGCAATCCACCCTC
CTCAGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCACCACACCCAGCCCCAGTGTAGTCGTTTTTTCTTTTCTTTTTTATTCTATG
TTTTAATGAATTTACACGTTACCCAAATGTTCCCTAGTTTTTCTGCCTTCCAAGATCACTCTGGAAGAATATTTAAGAATATACCAAAT
AAGAATATGCAAGTCCTCCCCTAAGGGTGGCAGGAAGAACACCCCTCCCCCAGATGGTATTTAGCGCCTCTGGCTGGGAACGGCTTCCC
CATGCTCCTAGGTCAGGGTCCTCTCTTGGCATGACACTACCACCACAGTGCAGACCCACAACAGGGAGAAGGACGGCCACAGTCCCTCA
ATCCCCCTTTTCCAAGATGTGCACAGCCTGACTCCTAACTCCCCACCACTGACTCTAGGGGAAAAACAGCACAGGGCAGGAAACGATTT
TCCATGTCACCAACCTTTCTCTGAGGGAACCTACTGGCCACCTCCCTCTTAGGACCAGCCCATCGTCCACAACGTGGAAGTCCAGCTTC
CGTTCAAATCGGAGTTCTTTCTTCATGACATTTCTTTGCAAAGTCCCGGAACCCACAGCTCTGAGACTCTGGCTGTCCCCCAACCCACC
CCATCTTCCTTGTCCTCACCCCTGGTCAGGAGAAGCCAAAACATCAGTCAGCTTCCCAGTAATCAAGCCTGGCTTTCTCACCCAGGGCT
CGCCCCAGAACAACCACCGGCTTCTTTCAGTGTAGCCAAAAGGCTATTGGAGTCTTCTCAAATGAAAGAGATTTTATCAAAGGCTTGGA
GAAGAAAAGAAAAAGAGGATTATATAATAAAACGTAAAACAACAAACATATACACACAAACAAAAATAAACGTGAGATATGATTCTCCC
GGAGTGTTTAGAGCAGGAATGTTCTTGGGCATCTGCCTTCCCCCACCAGCACCCCCCACAAGGCAAGGCCAGTTCACCCTCAGTGCTCA
CTACTTTGCAGTGTTCATAGAATATTTGTAATAATTTTAGGCGGCTCCCTAAAATTTCTTTTCTTTTTCTTTCTTTCTTTAGAGTTGCG
TCCCTCTCGGTTGCCAGGCTGGAGTTCAGTGGCATGTTCATAGCTCACTGAAGCCTCAAATTCCTGGGTTCAAGTGACCCTCCTACCTC
AGCCCCATGAGGACCTGGGACTACAGGTATGCACCGCTATACCCGTCTATCTTTTATTTATTTATTTATTTAGAGACAGAGTCTAGCTC
TGTCACCCAGGCCAGAATGCAGTGACACGATCTCAGCTCACTGCAACTTCTGCCTCCCAGATTTAAGGGTTTCTCTTGCCTCAGCCTCC
CTACTAGCTGGGATTACAGGCTTGCACCACCTACGTCCGGCTAATTTTTGTATTTTTAGTAGAGATGTGGTTTCACCATGTTGGCCAGG
CAGGTCTCGAGCTCCTGACCTCAAGTGATCCACCCGGCGTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACTACGCCCAGCCT
ATTTTATTTTATAATTTTGTTTTAGACAAGGTCTAGCTCTGTTGCCTGGGCTGGAGTGTAGTGGTGCAATCACGATTCAGTGCGGCCCT

FIG.3A-9

GATCTCCTGGGTTCGAGTGAGCCTTAGCCTCCTGTTTAGCTGGTACTACAGGTGCATGCCACCACCTAGCTAATTTTTTTAAAATTTTTT
TGTAGAGACGGGGTCTCACCCTGGTGTCCAGGCTGGTCTCAAACTCCTGGGCTCCAGTGATGCTCCCACATTGGCGTCCCAAAGTGCTG
GGATTATAGGAGTGAACTACTGTGCCCAGTCTTTTTAAAAAATTTTCAAGAGATTGGGGTCTTGCTATATTGCCCAGGCTGGTCTCCAC
TCCTGGTGTTAAGCGATCCTCCCACCTCAGCCTCCTTGAGTAGCTGGGATGACATTACAGGCACACACTGCCACCACTGGCTCTAAAAC
TTCTTCTGTGCCATTTGTGCACTTCACCCAATTGCCTCTTTGTAGTAATTAATTAGGATCTAGGGTGAAAAAAAAGTCAACAGCTATAT
ATAGTCCTCAAAGTTTTGTACGTATCTGAGCAGTCATCAGTTGCACAGTGCAGAGGGATGAACTGCCGTCCCGCCACCTAAAAAGCATT
AGTGACCATCAGGGAACCGTCAGATGCATGCCAGACTAAAGCAGAGTGAGGCTGTGCTGGGTGCTCTGTCTGTGGCTGCCCGTGCTCTC
ACTTCCCTGTCTTGCTCTGTGCCTTTGGGAGGTTGACCCTGAGTTGGCATCTCAGGGTCTCAGTCTGCTGGTTTCCTGSGTTCCCCTTG
AAGGCTACTGCTCCCACAAGGCAACCACGGTCCCCGCTCTGGCTCTCACTGAGCTCCAGAATCATTGTTTCCTCCCCTTACCCAAGTGA
GAATAATTATGTTTTATTCCAGAACCCTGACAAATGAAGAGGCCTAAAAACCCCCTAGGTATTATCCGATCTTGGTGATCAGGGAGGTG
TTTGTTTTGTTTTTTAATGCAGACACATAGTTTTAAAAATTATTCACTTCATCTACTGTAAGAAAAGTCATATTAATTCACAATTTTGA
TTAAAACAAACAAACAAACAAACAACTTCTGTGACATTTTGGCTAACAAGTGGTTCAATATTAAAGCTTTGTCCACCAGGTGCAGTGGC
TCATGCCTGTAGTCTCAGTGCTTTAGGAGGCTGAGGTGGGAGGATCACTTGAGGCCAGGAGGTCGAGGCTGCAGTGAACCATGATCTCA
CTACTACACTCCAGCCTGGGCAACAGAGTGAGACTCTGTCTCTAACAAACAAACAAACAAATAAGTATAGTTCTTTCAAGCATGGCAGA
CAATCTGTCTCCTTTGGCCTGGGTCTCTCACTGCCTTTTAGATAAAAATCTGGCAATAACCAAAGAGTTTTCATAAGGCCTGTTGATCT
ATTTATAAGACATGCATATAATTTACTTGACCATTATAATACCATTATAATAATCTAAATCTATTTTCTTTATCGTCCAATAATCCACA
GAGTCAGCACACAAGGATTCTTTTTTCCATATATAGGCTGAGTATTCCTTATCTTACATGCGTGACGCCAAAGTGTTTCAGGTTCTGGA
TGTTTTGGGATTTTGAAATATTTGCATATACACAATGAGATATCTTGGGGATAGAACCTACATCTAAACACAAAATTCATTTATGTTTC
ATATACACCTTATACACGTAGCCTGAAGGTAAATTTACACAATATTTTTAATAATTTTCCACATAAAACAAAGTTTGTATACATTGAAC
CATCAGGAAGCAAGGTGTCCCTGTCTCAGCCACCCACAAGGACACTCTGTAGTTGTCTTTCATTCCTGATTCCGAATTTATACGCTACT
GACAAGCAATCATTTTCTTACACTTATTCACACAAGAGCACTTAGTAAAAAATATGACATATATATCTGGCATGCTCAGAAAAGCTATT
TTGCAGCAGAAAGGAGCTGGGAGGGTCCTTTTTTTCCCTTGGGGACACGGAATAAATTGTGTATTATGTGCCTGCATTTTGACTGTGAC
CCCATCACATGAGGTTAAGTGTAGAATTTTCCACTTGTCTCTCTGTGCTTAAAAAGTTTAGATTGGCCAGGCATGGTGGCTCATGGCTG
CAATCCCATCACTTTAGGAGGCCAAAGCAGGTGGGTCATTTGAGGTCAGGAGTCAAAACCAGCCTGGCCAACATGGTGAAACCCTGTCT
CTACTAAAAATAAAAAAGTTAGCCTGGCATGTTGGTGCATGCTTGTAATCCCAGCTACTCGGGAGGCCGAGGCAGGAGAATCTCTTGAA
CCTGGGAGGCAGAGGTTGCAGAGAGCAGAGATCACTCCATTGCACTCCAGCCTGGGTGACAAAGCGAGACTCTGTCTCAAAAAAAAAAA
AAAAAAAAGGTTAGATTTTGGAGCATTTTGGATTTTGGATTTTGCATTAAGTGTGTTCAAGCTGAAAAGAAAATCCGATTTGCTCAGGA
CAAACTTAACAAAACAAGTGAGATATTCCAATACTATATATATGCTCCTGTTTATATTTCCTTAATTAATTTGGACTTGGAACAACTTG
GCCAATTATGGATTAGAGGATGAGACTTAAATGTTACTGTACAAGGGATAGAACGATTCATTCCTCTATGTTATCAAATACTTATGGTA

FIG.3A-10

TTTTMCCCATCCTGCTGTCATGCAGATCCAAGAACCAAATTAAAACACATTTGCCGGGGTCATAATAATGTGGCCAGAATTTAAAGAAA
AACTTGATTTTTAATTATGTATGATTTTGCTTGTTTAGTCTACCGATTTCTATTTGCTTTAGCTTACTCAAAAATAAAGCGCGGCACTT
CGAAGACTCAATAGTCTTCCATTCATGTGGGCCTTTATAATGCACGGGCCCAGATGCAATACATCTGGCGGTCTGCTTGGGTTGGCCAC
TGGATTGAAGGAGGCAGAGAAGTCTGGGATGATTCCCAAATGTCTGGATCTGGTGACAGGGAGATATGGCAGGGCGAGCTTAGGGGAAA
AAGCTGGGTTAGGAACTGTTGAAACTGAAATCCCTGAGGSYTKTGCCGACAGAGAGACAGCCGGTAGAAGGTTGTCTTTGCCTGTCTGT
GGTTCCAGGTAACTTCATCGAAAGAGAGTTTCAGGCAGTAGAAATAAGAGCACCCAGGACAAAGCCCCAGGGAAGAGAAACATCTGACG
GAGGACAGAGGAAGAAGGGTCAGGAATGAGACTGAGCAGGTGTCATGTGTCTGACACCAGAGCCTGACACATAGTACGTAGTAGACACT
CAGCAAATACCGTAACAGAGATGAATCCAAGGCTGGGGGAGGTGGCTCACGCCTGTAATCCCCACACCTTGAGAGGCCTAAGTGGGAGG
ATCTCTTGAGTCCAGGAGTTCGAGACCAGCCTGGGAAACATGGTGAGACCTTGCCTCTAAAAAAATAAAAATTAAACATTAAAAAAAGA
GATGAATGCATAACCTGGCTGCTGGAGCCAACATGGGTTGGGTGAGCCCACTCTTACCAGCAGCTAATCAAAAATTTGCCTGGAATTCT
GAGGCTCCTGTCCTACGTCTTGGCTGCTCCTCCCAGATCACCTTCTGGCCGGTCCCAAGTCCACTTCCCGTGCTCCTTGCTCCCTTCCT
CCTGGTCTCCCTCACACTTTCCTTTCCTACTCCCCTTCCCTCTGTGGCCCTGGCTCAGCCCAGCACAGGGAGAGCCCTGTGCCACCTAT
TACAGCTCACCTGCACCTTTGCATCTTTCAGAAAGGAGCACCTACAAGATAACCCACCCCCCACCTTTTTTTTTTTTTTTTTAGTAGTA
CAGATTGCCTCTCATAGCATAATTGGGCTTCATTATTATCCTTAAAGACCCTCTTTCTGTGGCGGATTGGGATGGATAAAATAAAGAAG
ATCGAGAGGTTGAAGAACCCATCCTGTTTTGCCAGTGAGAAGGGGATAGAATTAAAAGGATTAGGAGGGCTCAGGCATGGTGGCTCCAG
NGTGTCATCCCAGCTACTCAGGAGGCTGAGGCGGGAGGATCACTTGAGCCCAGGAGTTGGAGACTATAGAGCACTATGATTACACCTGT
GAATAGCCACTGCACTCTAGCCTGGGCAACATATCAAGACCCTGTTTCTAGGGACAAAAATATNNTTTAATAAATTTAAAAATTAAGGG
AAAGGTAACCACATCCTGCTACAAANAAAAGAAGNTGGAGAGGTANGANGAGGACCAAGAGCTAATGGCATCATTTACACAAAAAGAGA
TGCTTTAAAATCAGTTGCTCATCCAATTCCACAAGGACAATAAGTAAGAAAGAGGATAGAAAGTCACCGGTGGATTTGGTCATCATTGG
CTTCTTGATGACTTTAGCAACAAAAATTCTTGTTGGTAGTGAGAGTTAGACCCTGGTGGACTGGGTAGGGGGTTCCTGGATCATGAGCA
AAGGCCTGTGCCAGCCAATGGCCCCCACTACACTCTGCCCCGGCCTTTCTCATCTCAAAAAATGGCATCCCCCATCCAAAGCTCAAGTC
AAGAATCCAGCAGCCACCTTTGATTCTGCACTTCCCCTCACCTCACAGTCCAGTCCCATCTCCAAAATAAGTTCCAAATYTCACCACTT
CTCATTCTCCAAAGAGGMACMATTATCTCTTTCCTGGTGATTAAAACAGCTTCCTAACTGGSTTCCCTTCTACCTTGCTTTCCCATAGT
CCATTCTTCTCAGGACAACAACAGTGGCCTTTTAAAACCAGTGCATTATTGTTGCCCTTTGGGAAATCCTCCACAATTATCCAGTCTTG
CTTCAAAAAATGTATGTATTTCTGACTTTTTACCCTGCCCTACTTACAGGATATGCACATTTCTGATCTCCAGCCAATATCACACTTCT
TCTCTCACTGCACTCTGCCACACTTGGCCAAGTTTGTTCCCACTCCTCTTGCACTTGCTCTCAGATCTCAGAAGAGGCGTGCTCCTTGT
CTTTCAGGCCAGCCGGCTTCACACATGTGCCACGTGCGCCCCTCGCTCAGAAGGGATCTGTACTCGGTTTGGATCTATTGTTGCCATCT
TGAAACTCTTAATACTCTTTGAACACGGGGCCCGTATTTTCATTTTGCACTGGGTCCTGAAAATTGTGTAGCTGGCTCTACTTTCAGGG
ATTGTATCAGAAGTCTCCTCCTCAAAGAGGCCTTCCTCGGCCACTTATCCTCAAGTAGCTCCTCCCCTTCTAAGTTACTGGCTATCCCA

FIG.3A-11

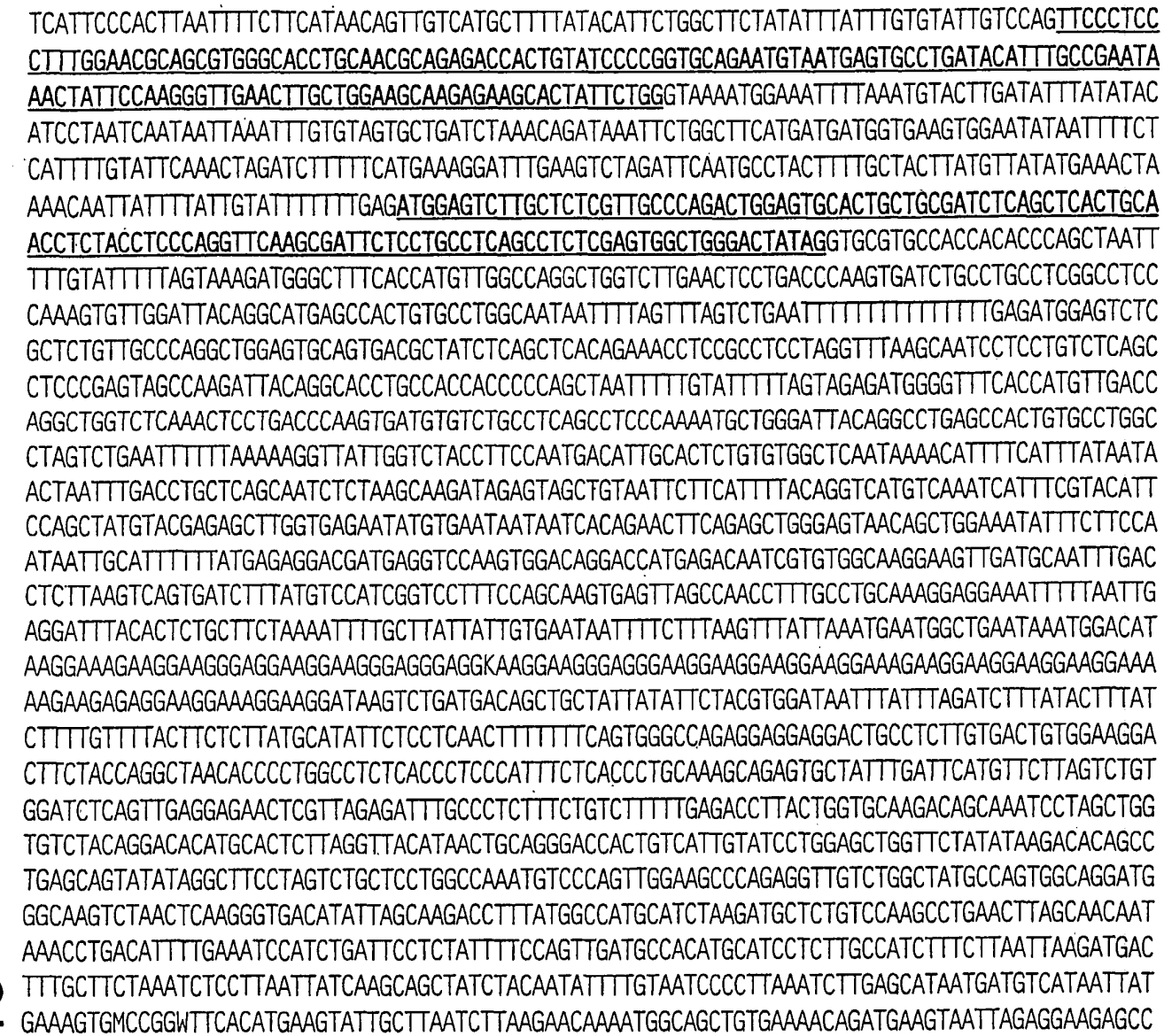


FIG.3A-12

TTTTTGGAAGCTTCGAGATATTTTCAAAGTAATTAGTACTAGTTAGCAATAAAGTTCTGTTCTGAGAAATTGCTCTTAAAGGAGGAACA
TGGATTAAAGAAAAAAATCTGCTACTAGGAAGTAAGCCATCTTCCTATGTGTGTGATTGGTTTTGCTTCCTGAAAACTGGTTCCGTTTT
CAACAAAATTTGGGTCTGTTGAAAAAGAACACGCAGATGCCAGCCTTGATGTCAAACGGGCCCAAACTTGGACAGTGGTAAACTAATGA
GCAATGGTGCACAGAGTCAGGGTAAAAGCTGGACAATTTCCTATGACCAACTTTTCCAGGACTCTGCTCTGCTCTTCCTGAGAAAAATA
CCCAAAGTGCTGCCTCTTCCATTGGCCCAACCATGCATCTTTCAGGATAGGMCACATCTGTTTATAGGTGTGGATTGTAGTTGCTCATA
AGTGACATTAGGCTGTTTAAAATAATAATAGTTCGAGTTTTGCTATGAGCTGATCTGTTTTCCAAGAGAGCTAAGAGTTTTCCAGCTAA
AAGAGGGAATTAGTGGGTAATCAAGGCAGCTGACATGGGGTGTGGCTGGGCCTTGAATGTGTGTCACTCTCTGTGCCCAGGCAGAGCAA
AGATAAACTCCAGACTGCATGTTGCTCAGAGACCAGGACCAACGTCATAGGGCGCCTAAAAGGCAGGTGGCCCAGTTCAGAATTGTCAA
GGTCTGACCTGCTTGGACAAGTGCTGAGTACATAGTAAGGATGGATTGGCTAGTCTCTCAAAACTTGCAAACAGGGCGCAGGTGATCTT
GAGATTTCAGGTGCCGGAGAGACCCATCGTGTAGATTCCAGAGTTGGCTATCATGACTAACAGCTGTCTAAGTTGTTTTTAAATGAATC
ATTAAGGGCTACATTTTCAGTTCAGCTAATCAAGTAGCAAATTACGGTGGGTCTAAAATACTTATCTATTGCATTATGTATATGCTAGA
CTTTATCACTTTAGTTGGTTATATCGCTTCATATACTAACAGTCAAAAAATGCCAAACGAGAAAACAAACAAACAAAAATGCCACATGA
CTGTGTAAATACACTTTTCAAACTGTTTTATCTAAGAGTTTACTCACTTTCACATTGTGGCTTATAGTATTTTCAATCTAAGAGACTAA
TTTTGCTTACATAGGAAACTACATATTTTAAATTGAAAATTAAAAAAATATTTTTAAGGTTTTAATGAGTCCTATCAAAACACATTTGT
ATATAGGAAGGTAGCCCAAGGTCACTGTTGCCAATTGTGTACACAGCCTGCCCTMTAGTGTTTTCTTCTAAACAGCACCAAATTTTAGA
TCATAGTTGTAAATCTCAAAATGTTGGGTTAATAGGATTAAACACTGTGTCATCAAATTGATAGGACACAGCTAAATCCCTGACACGGA
TGAAAATTAAAGCAGAGAAAAACGAAGGTCCTTCCAGAAGCTGGTGGCAACTTCACTGGGGAGATATTGCAAAGTTAGTGGTAAATACA
CTATATTAAAAAGTTTTGTTTTGTAAATAGAGTAATGATAGAAGAAGAGTTAGTTGAAATGATGTATGTAAAATGTGATAACTGCATAA
TTACTAGTACAGTTGCTAGTTTACGACTGTATTAAAAAGACATTCCAAATGTTGATCAAATAATGGAGGTTTCTGTGGTTGTTTTCTTT
TTAAAATAGTAAATATACGTAAAGCAGATAAATATCCCCTTTGTGGGAGTTAAAATAATCTAACTTATTTTATAGTTTTAACTTTATTA
AAGCATACGACTATTCTAACTTATTTAACTTTTCTTAGTAAAGTTTTAACCTCTGTATTTAGAATATTTGTAACTAATGTGTATCGAAT
TAAACTCAAAGGGAAATTCATTAACTGAGAAGAAAAAATTTTAACTGTGCACTATTCACATAGCATAATGGGTTTTATAAGGAGTATGA
GAAAAATGTGTGTGGTTGGTTTTGCTTTCTTTAAAAATAATAGCGAACCACGTAGGTAAAAACTCACTTGAGAACATAGACTTTTGGAG
GGAAATGCCAGGTGTGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGGGGGCGGATCACCTGAGGTCAGTAGTTCGAGA
CCAGCCTGACCCACATGGAGAAACTCCATCTCTACTAAAAATACAAAATTAACCGGGCTTGGTGGCGCATGCCTATAATCCCAGCTACT
TGGGAAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGTGGAGGTTGCGGTGGGCCGAGATCACGCCATTGCACTCCAGCCTGGGCA
ACAAGAGCAAAACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAGAATTTTGGAGGGAAAAAAATCCCTCTAACAGATTCGAATTAATTCT
GTGTTTCGAGATGTTTACAAAATGAAGCTTGGACTCTGAGAGGATGTGATCTATCCTCTCCATTGCATTGAGTTTCAAGTACTTCACAT
GGCGGGCTTTTTTAACTGTCGTGAAGTTTAAACCAAATAGGGACTAGAATTTGTTTGTTTTTTTAACTTACATTTCAAGCTTCCTTATG

FIG.3A-13

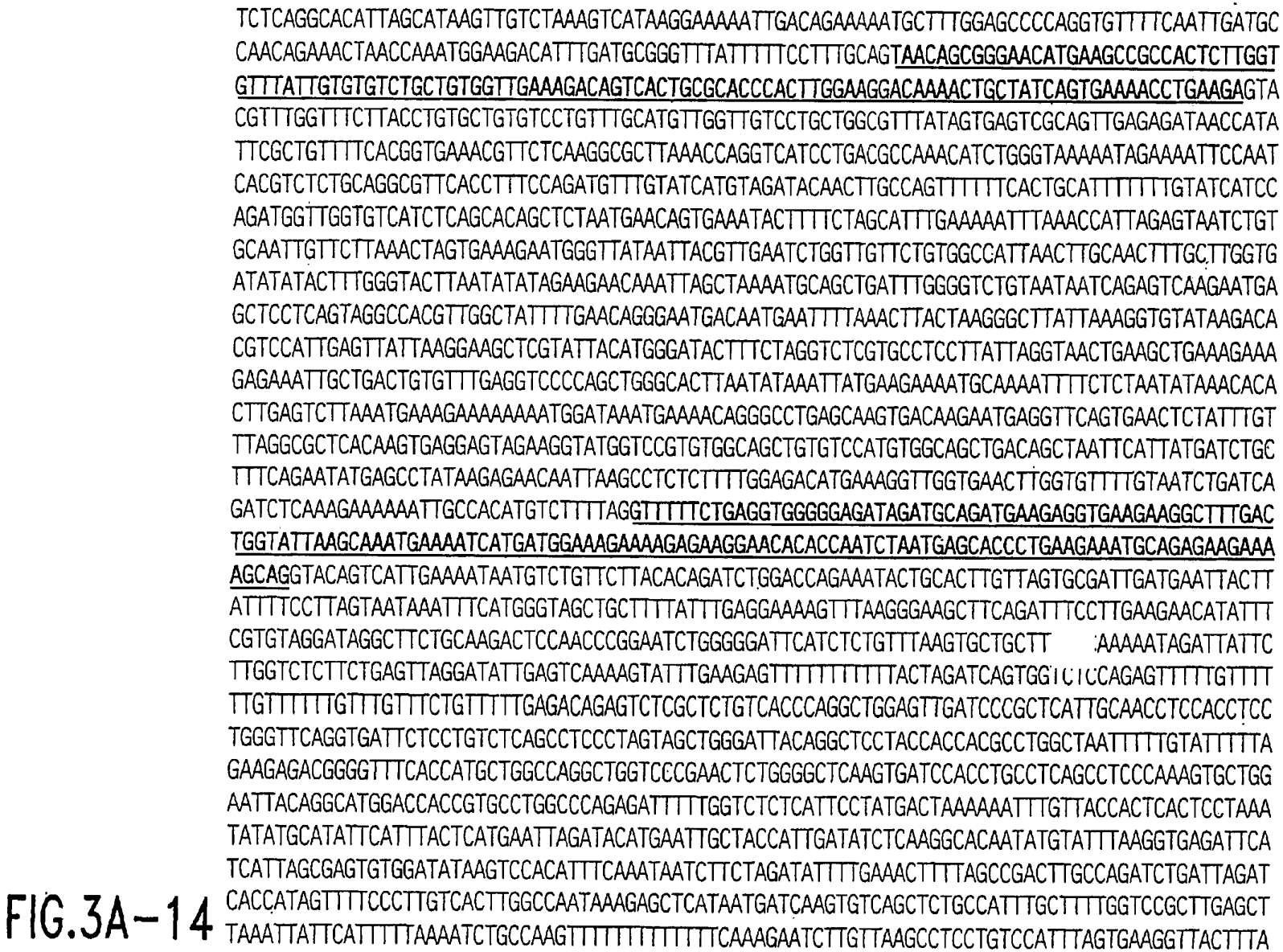

TCTCAGGCACATTAGCATAAGTTGTCTAAAGTCATAAGGAAAAATTGACAGAAAAATGCTTTGGAGCCCCAGGTGTTTTCAATTGATGC
CAACAGAAACTAACCAAATGGAAGACATTTGATGCGGGTTTATTTTTCCTTTGCAGTAACAGCGGGAACATGAAGCCGCCACTCTTGGT
GTTTATTGTGTGTCTGCTGTGGTTGAAAGACAGTCACTGCGCACCCACTTGGAAGGACAAAACTGCTATCAGTGAAAACCTGAAGAGTA
CGTTTGGTTTCTTACCTGTGCTGTGTCCTGTTTGCATGTTGGTTGTCCTGCTGGCGTTTATAGTGAGTCGCAGTTGAGAGATAACCATA
TTCGCTGTTTTCACGGTGAAACGTTCTCAAGGCGCTTAAACCAGGTCATCCTGACGCCAAACATCTGGGTAAAAATAGAAAATTCCAAT
CACGTCTCTGCAGGCGTTCACCTTTCCAGATGTTTGTATCATGTAGATACAACTTGCCAGTTTTTTCACTGCATTTTTTTGTATCATCC
AGATGGTTGGTGTCATCTCAGCACAGCTCTAATGAACAGTGAAATACTTTTCTAGCATTTGAAAAATTTAAACCATTAGAGTAATCTGT
GCAATTGTTCTTAAACTAGTGAAAGAATGGGTTATAATTACGTTGAATCTGGTTGTTCTGTGGCCATTAACTTGCAACTTTGCTTGGTG
ATATATACTTTGGGTACTTAATATATAGAAGAACAAATTAGCTAAAATGCAGCTGATTTGGGGTCTGTAATAATCAGAGTCAAGAATGA
GCTCCTCAGTAGGCCACGTTGGCTATTTTGAACAGGGAATGACAATGAATTTTAAACTTACTAAGGGCTTATTAAAGGTGTATAAGACA
CGTCCATTGAGTTATTAAGGAAGCTCGTATTACATGGGATACTTTCTAGGTCTCGTGCCTCCTTATTAGGTAACTGAAGCTGAAAGAAA
GAGAAATTGCTGACTGTGTTTGAGGTCCCCAGCTGGGCACTTAATATAAATTATGAAGAAAATGCAAAATTTTCTCTAATATAAACACA
CTTGAGTCTTAAATGAAAGAAAAAAAATGGATAAATGAAAACAGGGCCTGAGCAAGTGACAAGAATGAGGTTCAGTGAACTCTATTTGT
TTAGGCGCTCACAAGTGAGGAGTAGAAGGTATGGTCCGTGTGGCAGCTGTGTCCATGTGGCAGCTGACAGCTAATTCATTATGATCTGC
TTTCAGAATATGAGCCTATAAGAGAACAATTAAGCCTCTCTTTTGGAGACATGAAAGGTTGGTGAACTTGGTGTTTTGTAATCTGATCA
GATCTCAAAGAAAAAATTGCCACATGTCTTTTAGGTTTTTCTGAGGTGGGGGAGATAGATGCAGATGAAGAGGTGAAGAAGGCTTTGAC
TGGTATTAAGCAAATGAAAATCATGATGGAAAGAAAAGAGAAGGAACACACCAATCTAATGAGCACCCTGAAGAAATGCAGAGAAGAAA
AGCAGGTACAGTCATTGAAAATAATGTCTGTTCTTACACAGATCTGGACCAGAAATACTGCACTTGTTAGTGCGATTGATGAATTACTT
ATTTTCCTTAGTAATAAATTTCATGGGTAGCTGCTTTTATTTGAGGAAAAGTTTAAGGGAAGCTTCAGATTTCCTTGAAGAACATATTT
CGTGTAGGATAGGCTTCTGCAAGACTCCAACCCGGAATCTGGGGGATTCATCTCTGTTTAAGTGCTGCTT AAAAATAGATTATTC
TTGGTCTCTTCTGAGTTAGGATATTGAGTCAAAAGTATTTGAAGAGTTTTTTTTTTTACTAGATCAGTGGTCTCCAGAGTTTTTGTTTT
TTGTTTTTTGTTTGTTTCTGTTTTTGAGACAGAGTCTCGCTCTGTCACCCAGGCTGGAGTTGATCCCGCTCATTGCAACCTCCACCTCC
TGGGTTCAGGTGATTCTCCTGTCTCAGCCTCCCTAGTAGCTGGGATTACAGGCTCCTACCACCACGCCTGGCTAATTTTTGTATTTTTA
GAAGAGACGGGGTTTCACCATGCTGGCCAGGCTGGTCCCGAACTCTGGGGCTCAAGTGATCCACCTGCCTCAGCCTCCCAAAGTGCTGG
AATTACAGGCATGGACCACCGTGCCTGGCCCAGAGATTTTTGGTCTCTCATTCCTATGACTAAAAAATTTGTTACCACTCACTCCTAAA
TATATGCATATTCATTTACTCATGAATTAGATACATGAATTGCTACCATTGATATCTCAAGGCACAATATGTATTTAAGGTGAGATTCA
TCATTAGCGAGTGTGGATATAAGTCCACATTTCAAATAATCTTCTAGATATTTTGAAACTTTTAGCCGACTTGCCAGATCTGATTAGAT
CACCATAGTTTTCCCTTGTCACTTGGCCAATAAAGAGCTCATAATGATCAAGTGTCAGCTCTGCCATTTGCTTTTGGTCCGCTTGAGCT
TAAATTATTCATTTTTAAAATCTGCCAAGTTTTTTTTTTTTTCAAAGAATCTTGTTAAGCCTCCTGTCCATTTAGTGAAGGTTACTTTA

FIG.3A-14

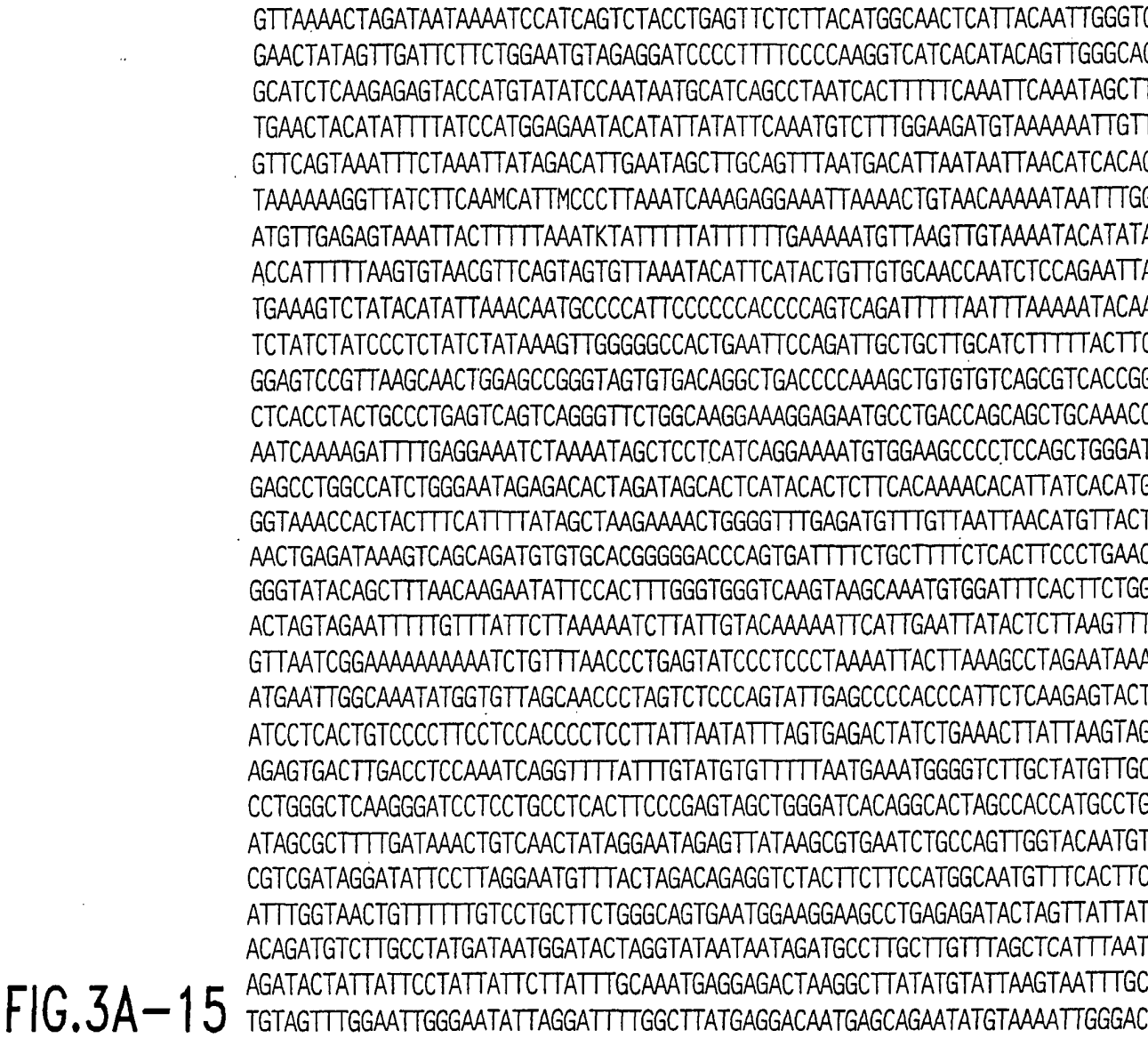

GTTAAAACTAGATAATAAAATCCATCAGTCTACCTGAGTTCTCTTACATGGCAACTCATTACAATTGGGTGCATGTGAACAGAGCAAGG
GAACTATAGTTGATTCTTCTGGAATGTAGAGGATCCCCTTTTCCCCAAGGTCATCACATACAGTTGGGCACACACAGTATCTGACATAT
GCATCTCAAGAGAGTACCATGTATATCCAATAATGCATCAGCCTAATCACTTTTTCAAATTCAAATAGCTTTATTTAACAGCTATAGCT
TGAACTACATATTTTATCCATGGAGAATACATATTATATTCAAATGTCTTTGGAAGATGTAAAAAATTGTTCATATGCCACAGTATAAA
GTTCAGTAAATTTCTAAATTATAGACATTGAATAGCTTGCAGTTTAATGACATTAATAATTAACATCACACTCAAAACAATGACTTTTT
TAAAAAAGGTTATCTTCAAMCATTMCCCTTAAATCAAAGAGGAAATTAAAACTGTAACAAAAATAATTTGGAAAATATTTTCAATTTTA
ATGTTGAGAGTAAATTACTTTTTAAATKTATTTTTATTTTTTGAAAAATGTTAAGTTGTAAAATACATATAACAAAATTTACCATCATA
ACCATTTTTAAGTGTAACGTTCAGTAGTGTTAAATACATTCATACTGTTGTGCAACCAATCTCCAGAATTATTTTCATCTTGCAAAAAC
TGAAAGTCTATACATATTAAACAATGCCCCATTCCCCCCACCCCAGTCAGATTTTTAATTTAAAAATACAAGTGGAAGTTCTAATATTT
TCTATCTATCCCTCTATCTATAAAGTTGGGGGCCACTGAATTCCAGATTGCTGCTTGCATCTTTTTACTTCTGAGCATCATGGCCTCTG
GGAGTCCGTTAAGCAACTGGAGCCGGGTAGTGTGACAGGCTGACCCCAAAGCTGTGTGTCAGCGTCACCGGACTGGTTGATGTTGCAGC
CTCACCTACTGCCCTGAGTCAGTCAGGGTTCTGGCAAGGAAAGGAGAATGCCTGACCAGCAGCTGCAAACCCTTCTCCCTTTTGGCAGC
AATCAAAAGATTTTGAGGAAATCTAAAATAGCTCCTCATCAGGAAAATGTGGAAGCCCCTCCAGCTGGGATCTTCCCTGGTGGGCTTGT
GAGCCTGGCCATCTGGGAATAGAGACACTAGATAGCACTCATACACTCTTCACAAAACACATTATCACATGGAATGTTTTGAACATCTG
GGTAAACCACTACTTTCATTTTATAGCTAAGAAAACTGGGGTTTGAGATGTTTGTTAATTAACATGTTACTCCAACACTGTAATGAATG
AACTGAGATAAAGTCAGCAGATGTGTGCACGGGGGACCCAGTGATTTTCTGCTTTTCTCACTTCCCTGAACCTCCTGGCAAGGAGGACA
GGGTATACAGCTTTAACAAGAATATTCCACTTTGGGTGGGTCAAGTAAGCAAATGTGGATTTCACTTCTGGCCCTGAAGAATCCAAGCA
ACTAGTAGAATTTTTGTTTATTCTTAAAAATCTTATTGTACAAAAATTCATTGAATTATACTCTTAAGTTTGAGGCACTCAATTAGAAA
GTTAATCGGAAAAAAAAAATCTGTTTAACCCTGAGTATCCCTCCCTAAAATTACTTAAAGCCTAGAATAAAGGTCAGTTTAGACAAATT
ATGAATTGGCAAATATGGTGTTAGCAACCCTAGTCTCCCAGTATTGAGCCCCACCCATTCTCAAGAGTACTGCTCAGTGGTGACCCAGC
ATCCTCACTGTCCCCTTCCTCCACCCCTCCTTATTAATATTTAGTGAGACTATCTGAAACTTATTAAGTAGGAAACCCTAGAGAAGGTT
AGAGTGACTTGACCTCCAAATCAGGTTTTATTTGTATGTGTTTTTAATGAAATGGGGTCTTGCTATGTTGCTCAGGCTGGTCTTGAACT
CCTGGGCTCAAGGGATCCTCCTGCCTCACTTCCCGAGTAGCTGGGATCACAGGCACTAGCCACCATGCCTGGCTCAATGCCAGGTTAAT
ATAGCGCTTTTGATAAACTGTCAACTATAGGAATAGAGTTATAAGCGTGAATCTGCCAGTTGGTACAATGTCTAGCAGGAAACGGAAGG
CGTCGATAGGATATTCCTTAGGAATGTTTACTAGACAGAGGTCTACTTCTTCCATGGCAATGTTTCACTTCCAAAACTTGGGACCTGTG
ATTTGGTAACTGTTTTTTGTCCTGCTTCTGGGCAGTGAATGGAAGGAAGCCTGAGAGATACTAGTTATTATACTGGACTAGTTATAATA
ACAGATGTCTTGCCTATGATAATGGATACTAGGTATAATAATAGATGCCTTGCTTGTTTAGCTCATTTAATGCAAAGACCTTGAGAAGT
AGATACTATTATTCCTATTATTCTTATTTGCAAATGAGGAGACTAAGGCTTATATGTATTAAGTAATTTGCCCAAGGGTACACAGCCAC
TGTAGTTTGGAATTGGGAATATTAGGATTTTGGCTTATGAGGACAATGAGCAGAATATGTAAAATTGGGACTGATTGAGAAAATCCTGG

FIG.3A-15

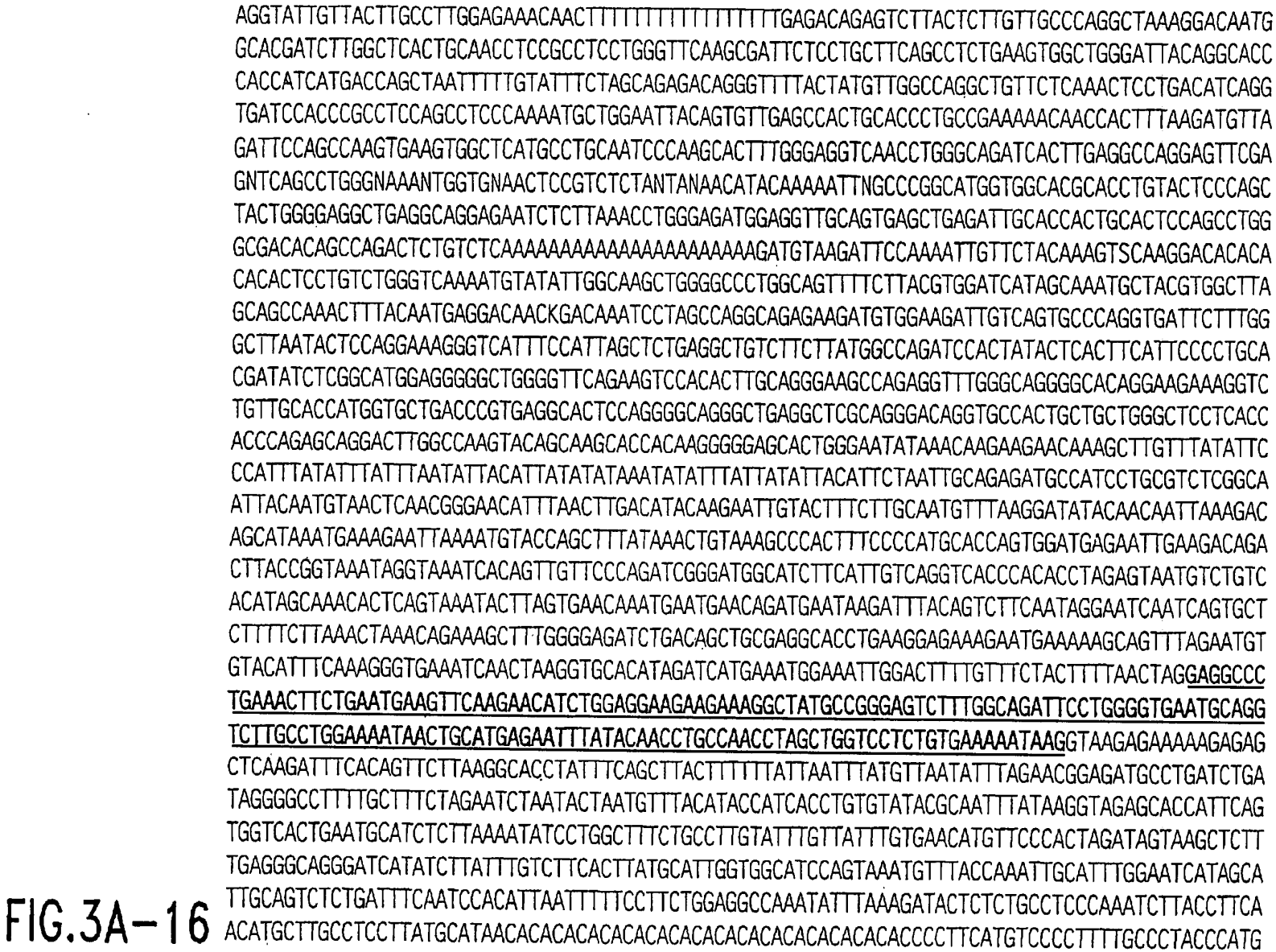

AGGTATTGTTACTTGCCTTGGAGAAACAACTTTTTTTTTTTTTTTTTTGAGACAGAGTCTTACTCTTGTTGCCCAGGCTAAAGGACAATG
GCACGATCTTGGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGCGATTCTCCTGCTTCAGCCTCTGAAGTGGCTGGGATTACAGGCACC
CACCATCATGACCAGCTAATTTTTGTATTTCTAGCAGAGACAGGGTTTTACTATGTTGGCCAGGCTGTTCTCAAACTCCTGACATCAGG
TGATCCACCCGCCTCCAGCCTCCCAAAATGCTGGAATTACAGTGTTGAGCCACTGCACCCTGCCGAAAAACAACCACTTTAAGATGTTA
GATTCCAGCCAAGTGAAGTGGCTCATGCCTGCAATCCCAAGCACTTTGGGAGGTCAACCTGGGCAGATCACTTGAGGCCAGGAGTTCGA
GNTCAGCCTGGGNAAANTGGTGNAACTCCGTCTCTANTANAACATACAAAAATTNGCCCGGCATGGTGGCACGCACCTGTACTCCCAGC
TACTGGGGAGGCTGAGGCAGGAGAATCTCTTAAACCTGGGAGATGGAGGTTGCAGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGG
GCGACACAGCCAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAGATGTAAGATTCCAAAATTGTTCTACAAAGTSCAAGGACACACA
CACACTCCTGTCTGGGTCAAAATGTATATTGGCAAGCTGGGGCCCTGGCAGTTTTCTTACGTGGATCATAGCAAATGCTACGTGGCTTA
GCAGCCAAACTTTACAATGAGGACAACKGACAAATCCTAGCCAGGCAGAGAAGATGTGGAAGATTGTCAGTGCCCAGGTGATTCTTTGG
GCTTAATACTCCAGGAAAGGGTCATTTCCATTAGCTCTGAGGCTGTCTTCTTATGGCCAGATCCACTATACTCACTTCATTCCCCTGCA
CGATATCTCGGCATGGAGGGGGCTGGGGTTCAGAAGTCCACACTTGCAGGGAAGCCAGAGGTTTGGGCAGGGGCACAGGAAGAAAGGTC
TGTTGCACCATGGTGCTGACCCGTGAGGCACTCCAGGGGCAGGGCTGAGGCTCGCAGGGACAGGTGCCACTGCTGCTGGGCTCCTCACC
ACCCAGAGCAGGACTTGGCCAAGTACAGCAAGCACCACAAGGGGGAGCACTGGGAATATAAACAAGAAGAACAAAGCTTGTTTATATTC
CCATTTATATTTATTTAATATTACATTATATATAAATATATTTATTATATTACATTCTAATTGCAGAGATGCCATCCTGCGTCTCGGCA
ATTACAATGTAACTCAACGGGAACATTTAACTTGACATACAAGAATTGTACTTTCTTGCAATGTTTAAGGATATACAACAATTAAAGAC
AGCATAAATGAAAGAATTAAAATGTACCAGCTTTATAAACTGTAAAGCCCACTTTCCCCATGCACCAGTGGATGAGAATTGAAGACAGA
CTTACCGGTAAATAGGTAAATCACAGTTGTTCCCAGATCGGGATGGCATCTTCATTGTCAGGTCACCCACACCTAGAGTAATGTCTGTC
ACATAGCAAACACTCAGTAAATACTTAGTGAACAAATGAATGAACAGATGAATAAGATTTACAGTCTTCAATAGGAATCAATCAGTGCT
CTTTTCTTAAACTAAACAGAAAGCTTTGGGGAGATCTGACAGCTGCGAGGCACCTGAAGGAGAAAGAATGAAAAAGCAGTTTAGAATGT
GTACATTTCAAAGGGTGAAATCAACTAAGGTGCACATAGATCATGAAATGGAAATTGGACTTTTGTTTCTACTTTTAACTAGGAGGCCC
TGAAACTTCTGAATGAAGTTCAAGAACATCTGGAGGAAGAAGAAAGGCTATGCCGGGAGTCTTTGGCAGATTCCTGGGGTGAATGCAGG
TCTTGCCTGGAAAATAACTGCATGAGAATTTATACAACCTGCCAACCTAGCTGGTCCTCTGTGAAAAATAAGGTAAGAGAAAAAGAGAG
CTCAAGATTTCACAGTTCTTAAGGCACCTATTTCAGCTTACTTTTTTATTAATTTATGTTAATATTTAGAACGGAGATGCCTGATCTGA
TAGGGGCCTTTTGCTTTCTAGAATCTAATACTAATGTTTACATACCATCACCTGTGTATACGCAATTTATAAGGTAGAGCACCATTCAG
TGGTCACTGAATGCATCTCTTAAAATATCCTGGCTTTCTGCCTTGTATTTGTTATTTGTGAACATGTTCCCACTAGATAGTAAGCTCTT
TGAGGGCAGGGATCATATCTTATTTGTCTTCACTTATGCATTGGTGGCATCCAGTAAATGTTTACCAAATTGCATTTGGAATCATAGCA
TTGCAGTCTCTGATTTCAATCCACATTAATTTTTCCTTCTGGAGGCCAAATATTTAAAGATACTCTCTGCCTCCCAAATCTTACCTTCA
ACATGCTTGCCTCCTTATGCATAACACACACACACACACACACACACACACACACACACACACCCCTTCATGTCCCCTTTTGCCCTACCCATG

FIG.3A-16

TATGTAGACTGGCATGTTTTCTTTTTTGTACCCTTTGGTTATCTTCTGAGCAGAGGGATCACAGAGGGTGGTGACCTGAATAGGATGAG
CTCTGCCCCACTAACGGCTCCAATTAAGCTAGATTTTTCTCCCCCTTCAAGAAGTGAGCTGAATACAAAATTGAGTGGAATTTCACGCT
CCATATTAGAGCACATACTAATTAGGGTATGCTCCTGGCTTGGCAATGCCATACTCAATTACAAAGGGAGCAACTACTAAGATAATGAA
TGCGCCAAGTTAATTTGCCTCCACTATTAATTGCATCTGCTCTATTTTTAGAGCTACTGTCGCCTGCTAATACACCAGAATATGGTGTA
ATCAGCACCAGCAGGAAGTCAGGAGATATGGGGACCATTCCCATCTGGGTCAGTTGTGTGATCTTATGAACATTTCTTGGGGCTTTAAA
GGTTTGTTTTTGTGGATGAAGAGTCAAGTAAACAGAAGCTGGTAGAGGGAGAGGCAGACAATCCACCCAAATTCTTTTCTTTATTTTTT
TTCATGAGACAGGGTCTGGCTCTTTTGCCCTGGCTAGAGGGCAGTGGTGCCATCTTGGCTTACTGCAGCCTCCACCTCCTGGGTTCAAG
TGATTCTCCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGCCCACCACCACGCCTAGCTAATTTTTTGTATTTTTAGTAGAGACAG
GGTTTTACCATGTTGGCCAGGCTGGTGACCTCAGGTGATCCACACACCTTGGCCTCCCAAAGTGAAAACTTGACCTTTTTAGGCTATTG
GTGGGCAATGTAAACCAGGAGAAATTTCAGATCCTGTTTCCATAGGCAAAGGCAAAGTCAGGTATAAGAGGGTTAAGAAATTATCTTAA
AGTTAATTGCCTCATACTAGCTTGCCCAGAATTATTATTGATTTGAAATGACTACTGTAAGTTGACTTTAAAATTTGCAATAAGAAATG
GTCCAGGGCCGGGTGCAGTGGCTCACCCCTGTTATCCCTAGCACTTTGGGAGGCCTAGGCATGTGGATTMCCTGAGCTCAGGAGTTCGA
GACCAGCCTGGGCAACACGGTAAAACCCTGTCTGTACTAAAATACAAAAAAAATTAGCCAGGCATGGCGGTGTGCAACTGTAATCCCAG
CTACTCGGAAGGCTGAGACAGAAGAATCACTTGAACCCAGGAGGCGGAGGTTGCAGTGAGCCGAGATGGTGCCATTGCACTCCAGCCTG
AGTGACAGAGCAAGACTCCATCTCAAATAAGAAAGAAAGAAAGAAAGAGAGAGAGAGAGAGAAAGAAAGAGAAAGAAAGAAAGAAAGAA
AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGRAAGRAAGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAAAAGAA
GAAAGAGAAAGAAAGAAAAGAAAAAGAGAAAGAAAGAGTTGAGAAAGAAAATAATTTTTTATTCCATTTCTGTCCCCTACTCTACTCCA
CAGATTGAACGGTTTTTCAGGAAGATATATCAATTTCTATTTCCTTTCCATGAAGATAATGAAAAAGATCTCCCCATCAGTGAAAAGCT
CATTGAGGAAGATGCACAATTGACCCAAATGGAGGATGTGTYMAGCCAGTTGACTGTGGATGTGAATTCTCTCTTTAACAGGAGTTTTA
ACGTCTTCAGACAGATGCAGCAAGAGTTTGACCAGACTTTTCAATCACATTTCATATCAGATACAGACCTAACTGAGCCTTACTTTTTT
CCAGCTTTCTCTAAAGAGCCGATGACAAAAGCAGATCTTGAGCAATGTTGGGACATTCCCAACTTCTTCCAGCTGTTTTGTAATTTCAG
TGTCTCTATTTATGAAAGTGTCAGTGAAACAATTACTAAGATGCTGAAGGCAATAGAAGATTTACCAAAACAAGACAAAGGCAAGTATT
AAAAGATTACTTTTACTTAGAGGTTTACACTAAAGTCAAGTTTTGTTTAGCTTCAGAAATGGTAGACATTTCTGAGTCACATTGTATAG
CGTTTCTTGAAGAGACAATTTATGGAAAATGTTTCAGAGCCTCTTAAAAGAAGCTTTGAAGTCTGCTAAACACTATCCCTCTTCCATCA
TCGTTGAGAACTGAACTCTTTCTAGAGCAAATTTTCAAAGCAGAAAGAAAAAATGCTAATAGGTTGAGAACTTGAAAAAAAAAAACCAG
TTCCCTCATTTATTATTTCTTTATTTATTTTATTTTGTGACGGAGTCTCACTCTGCCACCCAGCCTGGAGTACAGTGGTGTGATCTTGG
CTCACTGCAACCTCTGCCTCCCAGGTTCAAGCAATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGACTACAGTTGTGCACCACCACGCC
CAGCTAATTTTTTTGTATTTTTAGTAGAGACGGGGGTGTCAGTATCTTGGCCAAGCTGGTCTCAAACTCCCGACCTCAGGTGATCCACC
CGCCTTGGCCTCCCAAAGTGCTGGGATTGCAGGCGTGAGCCACCATGCATGGCCATTTCCCTCATTTATTAAAGCTCATGTAGATGCTC

FIG.3A-17

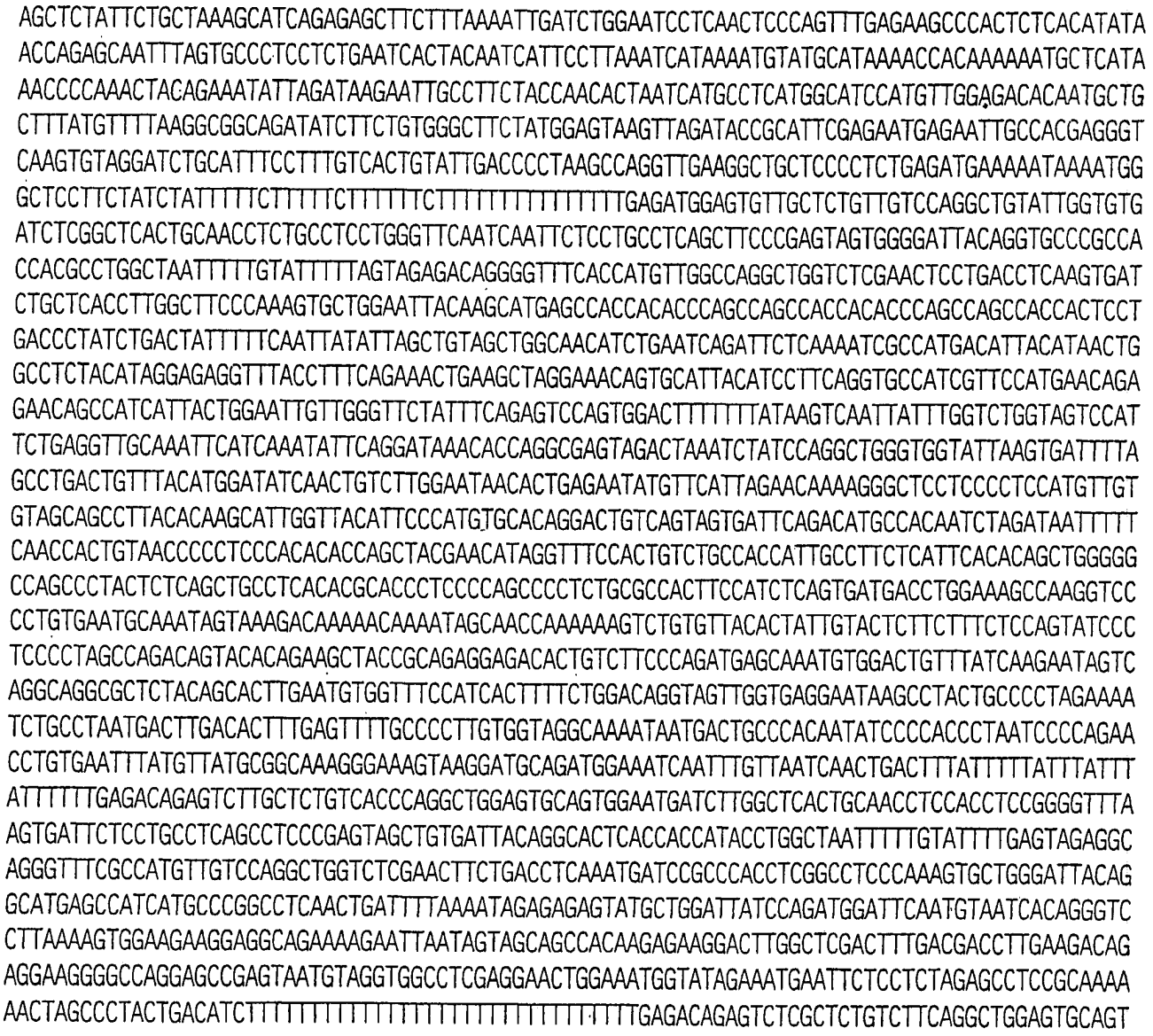

AGCTCTATTCTGCTAAAGCATCAGAGAGCTTCTTTAAAATTGATCTGGAATCCTCAACTCCCAGTTTGAGAAGCCCACTCTCACATATA
ACCAGAGCAATTTAGTGCCCTCCTCTGAATCACTACAATCATTCCTTAAATCATAAAATGTATGCATAAAACCACAAAAAATGCTCATA
AACCCCAAACTACAGAAATATTAGATAAGAATTGCCTTCTACCAACACTAATCATGCCTCATGGCATCCATGTTGGAGACACAATGCTG
CTTTATGTTTTAAGGCGGCAGATATCTTCTGTGGGCTTCTATGGAGTAAGTTAGATACCGCATTCGAGAATGAGAATTGCCACGAGGGT
CAAGTGTAGGATCTGCATTTCCTTTGTCACTGTATTGACCCCTAAGCCAGGTTGAAGGCTGCTCCCCTCTGAGATGAAAAATAAAATGG
GCTCCTTCTATCTATTTTTCTTTTTCTTTTTTCTTTTTTTTTTTTTTTTGAGATGGAGTGTTGCTCTGTTGTCCAGGCTGTATTGGTGTG
ATCTCGGCTCACTGCAACCTCTGCCTCCTGGGTTCAATCAATTCTCCTGCCTCAGCTTCCCGAGTAGTGGGGATTACAGGTGCCCGCCA
CCACGCCTGGCTAATTTTTGTATTTTTAGTAGAGACAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGAT
CTGCTCACCTTGGCTTCCCAAAGTGCTGGAATTACAAGCATGAGCCACCACACCCAGCCAGCCACCACACCCAGCCAGCCACCACTCCT
GACCCTATCTGACTATTTTTCAATTATATTAGCTGTAGCTGGCAACATCTGAATCAGATTCTCAAAATCGCCATGACATTACATAACTG
GCCTCTACATAGGAGAGGTTTACCTTTCAGAAACTGAAGCTAGGAAACAGTGCATTACATCCTTCAGGTGCCATCGTTCCATGAACAGA
GAACAGCCATCATTACTGGAATTGTTGGGTTCTATTTCAGAGTCCAGTGGACTTTTTTTATAAGTCAATTATTTGGTCTGGTAGTCCAT
TCTGAGGTTGCAAATTCATCAAATATTCAGGATAAACACCAGGCGAGTAGACTAAATCTATCCAGGCTGGGTGGTATTAAGTGATTTTA
GCCTGACTGTTTACATGGATATCAACTGTCTTGGAATAACACTGAGAATATGTTCATTAGAACAAAAGGGCTCCTCCCCTCCATGTTGT
GTAGCAGCCTTACACAAGCATTGGTTACATTCCCATGTGCACAGGACTGTCAGTAGTGATTCAGACATGCCACAATCTAGATAATTTTT
CAACCACTGTAACCCCCTCCCACACACCAGCTACGAACATAGGTTTCCACTGTCTGCCACCATTGCCTTCTCATTCACACAGCTGGGGG
CCAGCCCTACTCTCAGCTGCCTCACACGCACCCTCCCCAGCCCCTCTGCGCCACTTCCATCTCAGTGATGACCTGGAAAGCCAAGGTCC
CCTGTGAATGCAAATAGTAAAGACAAAAACAAAATAGCAACCAAAAAAGTCTGTGTTACACTATTGTACTCTTCTTTCTCCAGTATCCC
TCCCCTAGCCAGACAGTACACAGAAGCTACCGCAGAGGAGACACTGTCTTCCCAGATGAGCAAATGTGGACTGTTTATCAAGAATAGTC
AGGCAGGCGCTCTACAGCACTTGAATGTGGTTTCCATCACTTTTCTGGACAGGTAGTTGGTGAGGAATAAGCCTACTGCCCCTAGAAAA
TCTGCCTAATGACTTGACACTTTGAGTTTTGCCCCTTGTGGTAGGCAAAATAATGACTGCCCACAATATCCCCACCCTAATCCCCAGAA
CCTGTGAATTTATGTTATGCGGCAAAGGGAAAGTAAGGATGCAGATGGAAATCAATTTGTTAATCAACTGACTTTATTTTTATTTATTT
ATTTTTTGAGACAGAGTCTTGCTCTGTCACCCAGGCTGGAGTGCAGTGGAATGATCTTGGCTCACTGCAACCTCCACCTCCGGGGTTTA
AGTGATTCTCCTGCCTCAGCCTCCCGAGTAGCTGTGATTACAGGCACTCACCACCATACCTGGCTAATTTTTGTATTTTGAGTAGAGGC
AGGGTTTCGCCATGTTGTCCAGGCTGGTCTCGAACTTCTGACCTCAAATGATCCGCCCACCTCGGCCTCCCAAAGTGCTGGGATTACAG
GCATGAGCCATCATGCCCGGCCTCAACTGATTTTAAAATAGAGAGAGTATGCTGGATTATCCAGATGGATTCAATGTAATCACAGGGTC
CTTAAAAGTGGAAGAAGGAGGCAGAAAAGAATTAATAGTAGCAGCCACAAGAGAAGGACTTGGCTCGACTTTGACGACCTTGAAGACAG
AGGAAGGGGCCAGGAGCCGAGTAATGTAGGTGGCCTCGAGGAACTGGAAATGGTATAGAAATGAATTCTCCTCTAGAGCCTCCGCAAAA
AACTAGCCCTACTGACATCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGACAGAGTCTCGCTCTGTCTTCAGGCTGGAGTGCAGT

FIG.3A-18

GGTGCGATCTTGGCTCAGTACAACCTCCGCCTCCTAGGTTCAAGCGATTCTTCTGCCTCAGCCACCTGAGTAGCTGGGACTACAGGCAC
GTGCCACCACGCCCAGCTAATTTTTGCATTTTTTTTTTTTTTGAGACAGATGACATCTTGATTTTAGCCTAGGGAGACCCACTTCAGACT
TCTGACCTAAAAGACCAAACAATAATGAATTTGTGCTGTTTCAAGCCACTGAATCTGTGGTAGCTGTAGCAGAGCTAATAATAATAGTA
ACTGACCAACATTTACTGAGCAAGTTCCGTGTGGCAACCTTCATGGATGGGCCTTATTGGTCATGATTGTTTAAAGGGCCAAAATTAGA
AAAATAGCTAACACTGAATTATGAACACCAGGAAAAGGAGAGCGGAAATAAAAAGAATCAGAAATATCTTGATAATTAATGCTATTTTT
GTTGAGTATAGGTTCATTTTGTTCTCATATTTCTTTCCTACCTTGGTCTTTCTGGACCTCAGTTCCTGAATCTGTTGAAAGCGAATAGG
TCCAGGAAAGTAGCTCTTGGAATTATCTTCATTTGCCTTATGAATCCCTGGAAGGAACAGATGAGATTGAGTTCTACTGTAGCTTGACC
CGTGCGGGGGCCGGGAGACCTGGTTCTAATGCTGCCTTAGAGAGTGTTAGTTAACATTAATTTTCGCGTGGGAGAAACAGACAGGCAGG
TGGGAGAGTAGATGATTTAGCTCAGTGACTGCACTGGAAGTAGCTCCCTGGAAGGGTTCTGAGGTTCTGTCAAGGCTAGACTAAGCGAG
GTGATGGATTGTGCTGTGGCTGCAGGATGGGGAATTAGTGTCATATGGGCCTAGAATTTGTCATCCTTGGTGTACATACCAGGTATTAA
TCTAGATGCTAGAGATAAAATGATGATTATGACACAGCCTCTGACTTCCAGGAGCTCAGTCCAGAGAAAGGAAAACAGATTAGTGAACA
ATTACATCACCATATTGTGGGTAAAATGGCAGAAGAAGGTATGGAAGAATGACAAGATTAAAATGGCAAGACCAAGTCCCTTCCCTCAA
GAGGCTTACAGTCTAATGGAAAAGATAAGAAAGCAAACACTACATAAAGCAGGAATTAATTCTACACTGGAAATTCTCACAGGGGGCTA
TACAGGGCAAAGAAGAGGGTCCAGGAAAGCAGCTGGGAGAAACTGACTTTCTGGTCACCAAAGGGGATGGGTGCCTTACATGCCATTCT
ATCAAACAGTGCTTCACTGTTTTTAAACTATGGACTTTGCAATTTATCTCAAAATAAAACGTTTCATTTTTAAATGCTGAGGATTTAAT
ATGACAGAAAATCATCAGGTTGTAAATTAGTAATACATGTTTCCTAATGTCAAACACTCTATTGGGAACCGCCAATTTTCTGTTGGATA
GACTTCTCTTTTACACATTTTTATATGGATTGTTAATTCTCCTAGGGGAAAAAACTTCTCAAAACTTGATTGGCTTTAGATATTTTCCT
AAATCTTTGACCCCCTGTTCATAACAGTATATGCATCTCCACACACACATACTCGCACACATATGTGTGTATATATATGTGTGTGTGTG
TGTGTGTGTGTATATACATATATATGAGAAATGCAAAAAAAGAATAGTAATAAAATAACCACCTATCACCCACTTTAAGAAACAGACAT
TTCTAATATCTTTGAAACTTCTTCCCAATTATAGCTTTAAAAATTAATTATTAAAGAGTTTTTTAAAATACAGAAAAGTCCAAGAGAAA
AAGTGGTTCACAATCACCTATTTACTTAATCCTATTGACATCAGAAATACTAATGATATAAGACAAATGATTTTTAAAGTAATCAAATA
TATAAAAGAACAAAATAAATGAAAGCTGCCCTCTCCTACCTTATCAACTCCCTCTTCTAAAAGATAGTTATTAATAATTCTTCATGACT
CCTCCTAGAAAATAAAATTACATGCATTAATATATGTGTGTATATACTACTAATAAATTTCTAGTAATGAGATTCTTGGATTCAAGAGT
GTGCAATTTTTAATAGCTGTTCAGTTGTCCCAGGAAATTATTGCACCAACGTGCATTTCTGTGTCTAAATATAGGAAAAAGGGCCAGGG
GCGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCATTTGAGGTCAGGAGTTCAAGAAACCGGCCTGGAC
AACATGGCGCAACCCCATCTCTACTAAAAGTACAAAGATTAGCTGGGCTTGGTGGCTCTCACCTGTAATCCCAGCTACTTGGGAGCCTG
AGGCAGGAGAATCACTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCAAGATCCCGCCACTGCACTTTAGCCTGGGCAACAAGCAAGAC
TCTGTCTCAAAAATAAATAAATTAAATACATACATACATATAGGAAAAAGATTTTGAAAGCACTGGTAAGAAAAAGCTGCGGCATTGTC
TCCACTTCTTCAAAGTGCAAACTCTTATGACACTAACGTGTAAATGTTATGTTCCCTGTAGCTCCTGACCACGGAGGCCTGATTTCAAA

FIG.3A-19

GATGTTACCTGGGCAGGACAGAGGACTGTGTGGGGAACTTGACCAGAATTTGTCAAGATGTTTCAAATTTCATGAAAAATGCCAAAAAT
GTCAGGCTCACCTATCTGAAGGTAAATAATTGCTATTTTGTTTTTTATTCTACTTTAAGTTCTCAGGTACATTTTGTTATAAAGTTTCG
GTGCCACAAAAGAAATAGCACTCGAATATAAAATTTTCTTTTTAATTCTCAGCAAGGAAAGTTACTTCTATAGAAGGGTGCGCCCTTAC
AGATGGAGCAATGGTGAGCGTGCACTTGCCAAGGGAGGGGAAGGGGTTCTTAACCCTGACAATGCACGTGGCCCCTGCTGCTGTGTGGT
TCCCCTATTGGCTAGGGTTAGACCGCACAGGCTAGACTAATTCCCATTGGCTAATTTAAAGAGAGTGACGAGGTGAGTGGTCTGGAGGG
AAAAATGGTTATGACAGAGCATGTAATCGGAATGAATCAGGGCGGAGCGTGTAATCGGAATGAATCAGGGCGGAGCATGTAATCGGAAT
GAATCAGGGTGGAGCGTGTAATCGAAAAAGGTTGCTTTACGAGGAAATTAAGTTTAAAAGTAGAAGGCAAAGAATTGAACATACTGACA
TACTGATTCTTTGGAAAGAAATTTAGAACTCACATCTAACAATTTTTTAGGGTTTCTTTAGTATTCTGGACAGAGGACAAAATCTCATT
CTCACAAGCATAGTGGATTCATTTGCTTTCCTCCAAGCACTTTTTTGCAGGCTCATTTCCATCTGGGGGCGTTCAATGTAGGTTTATAA
ACTGGTGTTTTGTTTGTTTGTTTTATGAGACAGAGTCTTGCTCTGTTGCCCAGGCTGGNGTGGCACAATCTCGGCTCACTGCAACCTCC
ACCTCTCGGGTTCAAGCAATTCTCCTGCCTCAGCCTGCCAAGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTTTT
GTATTTTTAGTAGGGACGAGGGTTTCACCATATTGGCCAGGCTGGTCTCGAACTCCTGACCTTGTGATCCGCCCACCTCGGCCTCCCAA
AGTGCTGGGATTACAGGCATGAACCACCGTGCCTGGCCTGGTTTATAAACTTTTATTATTCCAAAGTATGTCATTCTTTCACTTTCTTT
AATTCCCTAATTGTTCTTGTGATTTTTTTTATGATTAATGACCAAACACTATTGTGTGCAAAAGAAAAACCTTGAGCAAATTAGCGCAA
CTCCTTCCTTCTTACCGCAAGCAAAAAGAACCCCTGCCCCCAACCATGAAAGAAACCTTTCATTCTGTAAATCAGTGTTTAGACAAGTG
AAATATTTTTTTGAAAGTGGCATTGGCTCTTTCCCATTGGTGGGTTAATGAACTAATTAGCATTTAAATAGGGAAAGTGGCTTCTCCTC
CCAAGCCCCAGGAATCCTTTTCCCTCCCTTTCTAGTTCCTTCCCCAGGAAGGAAATCATTCTCCCTTTCCTCCATCCCTCCCCTCATTC
CCTTTCCCTTCTCCAGACTAAAGTCACTCCTCCAACCCCACCAGGGCCAAATTACAACTTTTCTTACATAAAACAAGAGCTTTTGATTC
CTATGCTTCTGCATTTTATCTCACTAAAGCCCTAAGGGAAGGAAATTTTCAAAGTGTGACTAATGGCTTACAGTAGGAAATTGGAAGAT
ACAGAAGGGACAGAAATCAACATGTCAGTAAATTCTACAACACTAGCTAGAGATTTGGGGCAAGTCATTTATGCTGTCTAGGCCTCAGT
TGAGTAATTTGTAAATAAAGGACCCAAGATAATCTTTGGGTTCTAACAAAATTCTTCTGTAAAACAGTGGTCCCCAGCCTTCTGGCACC
AGGGACTAGATTCCTGGAAGACAATTTTTCCAAAGATGGTGGGGCAGGGGGCACGTTTGGGGATGATCATCAGGCATTATTCTCCTAAG
GAGCGCTCAACCTAGACCCTTTGCATGCACAGTTCACAATAGGGTTTGTGCTCCCGTGAGAATGGAATGCCTCCGCTGATCTGACAGCA
GGCGGGGCTCAGGCAGTCATGCTTGCTCACCTGCCGCTCACCTCCTGCTGTACAGCTCCGTTCCTAAGAGGCTACAGGCTGATATGGGT
CCGTGGCCCAGGGGTTGGGGACCCCTGCTATAAAGGAAGTTCAGAAAAATCAGATTATAATTCTGATTTTTATAAATCAGAATTTATAA
AATTCAGATTATAATTTACTACCAAGTAATAGCTCTTTTGCCCTTAACTTCCCACAGTGAAGACCACTGGAGTAATTTATATCAACGCA
AAGAACAAAAAGCATGGTCAGTGGAAACTCCTGCCCCTCCCTTGGCTTTCTCTCCTCAATCTAACAGTGAGCAAGTTGCAACAAATCGC
GCCGTTCAGAGAAAAGGGAGGATGGAATTGTTACAACCGTTTCTGTCGCCCAGGCTGGAGTGCAGTGGCGCGATCTTCGCTCACTGAAA
CCTCTACCTCCTGAGTTCAAGCGATTCTGCTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCACGCGCCACCATACCTGGCTGATTT

FIG.3A-20

FIG.3A-21

AAAATGTAAAGAATGTCCTAATGTGCTCCCATGCTGCTTAAAACTGTTATTATAAATTGCTTTTTATTATAAATATATAAAGAATGATG
TAATAGGCCAGCCATGGTGGCTCATCCCTGTAATTCCAGGTCTTTGGGAGGCTGAGGCAGGTGAATCACTTGAGGTTAGGAGTTTGAGA
CCAGCCTGGCCAACATGGTGAAACCCCGTCTCTACTAAAAATATAAAAATTAGCCAGGTGTGGTGGTACGCACCTGTAGTCTCAGCTAC
TCCGGAGGCTGAGGCAGGAGAATCGCTTGAAACCAGAAGCCGGAGGTTGCAGTGGGTCAAGATCAAGCAACTGCACTCCAGCCTAGGTG
ACAGAGCGAGACTTTGTCTCAGGAAAAAAAAAAAATTCTCAGTCACCTAGATTGAGAAATAGAACATTACCAAAACAGATAAAGCCCCA
CTGTGTTCCCATCCACATCACATTCACTTTATCTCCTCAAAAGGAAAGTGCTATTTTGAATTTAGTATTAATTATTTCCTTGCATTTCT
TCCTACTCATATCATGTGCCTATATACATATAATATATACAAATGCCGATATCATACATAGCAATGTTTTACATTTCGATTTTTGCATT
GTCAATGTAGAATTTTTAAACTTAAAAACATGCTTCATACAGCCGGGTGTGGTGGCTCATGCCTGTAATCCCAGCATTTTGGGAGGCCA
AGGCAGGCGGATCGACGAGGTCAGGAGTTCGAGACCAGCCTGACCAACATGGTGAAACCCCATCTCTATTAAAAATACAAAAAAAAATA
TTAGCTGGTCATGGTGGCGCGTGCCTGTAATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATTGTTTGAACCCAGGAGGCAGAGGTTG
CAGTGAGCCGAGATCGCACCATTGCACTCCAGCCTGGGTGACAGAGCGAGACTCCATCTCAAAAAAAAAAAAAAAAAAGCTTCATACAAA
CATGAAACGGGCACATGTCTGGCTGGGTGCGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCCAAGGCGGGCAATCACTTAAGG
CCAGGAGTTCGAGACCAGCCTGGTCAGCATGGTGAAACCCCGTCTCTACTAAAACTACAAAAATTAGCCAGGCATGGTGGCATGCGCCT
GTAGTCCCAGCTACTCGGGAGGCTGAGGCACAAGTATCACTTGATCCCAGGAAGCAGAGGTTGCAGTGAGCCAAGATTGTGTCACTGCA
CTCCTGCCTGGGTAACAGAGTGATACTCTGTCTCAAACAAACAAACAAAAAAAACAAAGAAAAGAAAAAGAAAAAAGAAATGGGCACAT
GTCAAATGTTAATTTGACTATGTAACTTATTAATGAAGGAACCAGCAGGGTGTTAGAGCTGGGTCAAAGAAGTATAAGAGAGACTGGAG
TGCTTACAGTCAAGCAGAGACAGAATGCTGAAAGGTTATGAAATTAGATATGTTAGTTAATATTCGAAAGGGCAACTAAACTGTAAATC
TTGCCATTATCTTTTCTATCAGACCAAAATAATTTACATCTCTACTAGACAAACATTTGCCACTTTTCAATCCATAATCTATGGGTAAT
TTCATGGAGTCTGGCCCTAATCAACAGTAAATAGTAAAGCCAACAAAGGATCTCTTCCCTAGACCTTGAAGTGATCTTTGGGTGGACCC
CTTAGACAATAATTTAGTATGACATTGAGAGGACACGCAAGCCTGGGCAGCATAGTGAGACCCGCCTCTACAAAAAAATTAAAAATTAG
CCGGGCATGGTGGTGTGAGCCTGTAGTCCTAGCTACTCAGGAGGCTAAGGTGGAAATACCACTTGAGCCCGGGAGTTCGAGGCTGTAGT
GAGCTATGATCATGCCATTGCACTCCAGCCTGGGTAACAGAGCGAGAACCTGTCTTGAAAAAAAAGAAAAGAAAAAAGAAAAAGAAACA
AAAGGAAATGCAGCCATTTTTTTTTTTGCCTTATTTCCAAGTTCTGGATAATTTTTCTTTTTTAACAATATAAATATTATCACTTATGTA
TTCTTTTGCAATATGGCTTTTCACTCAGTGTAGTTTGCAAGGGGTTAGCCATGTGAATGCATGCTGCTCTAGTTCATTAATTCACTGTT
GTATGTTGGTCTATGTAGGCATATCACAATWTATYCATTCCCTAGCTGAAGTACATTTGCTTTCAAGGTATTGCTATTATAAACAAATC
TCATACCTTTAATCAAATAATAATTTTGTCTCTTCAATCAGCTNTGATTTACTTTGTTCNAANACNAAGCACACAACTATAATTANAAT
TTCATTACTGATAAATATAAAATATTTTCCAAAACATCACAAATCTTTTNTNNTNCACTATTTACTATACACTTTNGGTCTNAATTTAA
AGCGGCTTCACTATATGTGGTTCTTTTCCTCTCTTCCCATACTAATTACTGGTACTGGACATATACATCCAAAATCAAATAGTARTGTC
CTTTTTAAGGGATAAATGGGATGTGATGTAGAAGGGGCATAGTAGGGACTTCATCTGTTTTGGCAAATTTTTTCTTAATATAGGTGGTA

FIG.3A-22

GGCATGTGGAATTTATAACAAAAGTTCTGTCTCCAGCCCAGTTTCTGTTACATAAAACCATATAATTAACAGTTAAACTGGATCTGGTT
TGACACAGATGTAGACGATATTAATAATTACTCCAGAACAACAGGCATAACTAAAAACTACCACAGGCAAAAGGGGAAAATAGAGAATG
TAAGGGCTGGGACTTAAGCCCATGTTGCCCACCTCCAAGTTTCATGGACTTTTTCCTTCTCCACATTACTTTCTTCTCTGCTAGACTGT
CCTGATGTACCTGCTCTGCACACAGAATTAGACGAGGCGATCAGGTTGGTCAATGTATCCAATCAGCAGTATGGCCAGATTCTCCAGAT
GACCCGGAAGCACTTGGAGGACACCGCCTATCTGGTGGAGAAGATGAGAGGGCAATTTGGCTGGGTGTCTGAACTGGCAAACCAGGCCC
CAGAAACAGAGATCATCTTTAATTCAATACAGGTAAAGGAGAGACCCAAGAGCAGATACGGAAATGACACGTGCATACCTTGATTTCAC
TGTTAATTTACTTATGAATTGTGTCTGAATTTGAAAACAAGCTGTAGGAGGTATTCATATTTCCATTGTGATTGCCTTCAGGCTGACTT
GATTTAACGTAGTTCATGGTCTTTAGAAAACAAGAAAGTCCATAAAGAAAATCAATTTAAAACACAAAATACTTTCTAATCTAGAAATG
GCTATTTCTGCTTAGAGTTATAGGGCTATAACTGATAGAGGTAACCTTGAAGAAATATGGCCAATGTAGGTTTTAGGAGAGAAGACTTA
CAAATAAAGCAATTTGAGTTCAAAATTTGACTCTGAAACTTACCAGCTGAGTAAGCTTGGGAAAGTACCTCAACCATTCTAGGCCTCAG
TGTTCCACCTGTAAAATGGTAACAATCATAGCTATCTTAACGTGTACACCTATAAAGTGATTAGTATAGATTTCTTATACAAAACAAGA
GCTCTGTAAATTATAGCTCTTATTAGTTGCTGACACAATAAAGCCACTGAGTTATCTTGAGAATTAAACATTTATATGTTACTCGTCAC
ATAAAAATACATTGCCAGCTGGGCGCAGTGGCTTATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGTGGGTGGATCACTTGAGGTCA
GGAGTTTGAGACCAGCCTGGCTAATGTGGCGAAACCCCGTCTCTACCAAAAACATAAAAAATTAGCCAAGTGTGATGGCACACACTTGT
AATCCCAGCTACTCAGGAGGCTGAGGCAGGAGAATCACTTGAACCCGGAAGGCAGAGGTTGCAGTGAGCTGAGATCGTGCCACTGCACT
CCAGCCTGGGCGACAGAAGGAGACTCTGTCTCAAAAAAAACAAAAATAAARACATATTGCCATCTTAAATTCCACCTATACCATGACTC
CCAGATTCAGTCAATAACTTTTTGCATAACATGCAAGTGACTTTTCTTCCTAAGACATCCCCCCTCCAACACACACACATTACCTTAAT
CTACAAATGCGCCAGGCTAGTGATTCCTGATGAGGCTGGTTTTGAGGGTTCCCAAAAAGACTTGGATACAAAAATTACTGGGCAGAGCA
ATTGAAGATGCAATATTCTGTGTGTAGTATGTTAGGTTATGTTGGTGCCCTATCCAGATCCCTGGGGATCCCTTTTACCAGCTCCCACT
GGTGCTGGTGCTGCTGCTAACTGCTTATCTCTGAAACTTTCTCCCAAAGATTGCCCTTGGAGCACTTATGCCCCAGAGCTTCCTGCAGG
ATCAGGCTGAGGCTAACAGTCATCTGAAGCCATATCCTTGCTTAGCTTCTTTCACTTCTCTAGTTTGCTTTCCTCATCCCCTTAAAAGT
TGCACCTGAGAGCATTCTTTATAAACCACTTCTGTCAGAATCTCAGGCACTGCTTCTAGGAAATTAGACTTATGGCATTCTATAATCCA
GCATTTCCCTCTTTTTTCAAACTACAAAGCTGTGGATCATGCCTGATTTGAGAAATAAGTTTAGAAAGTCACAGCAAGCTCATTAAAAA
ACAAAATTAAAAACCATACAAAAAATAGAATAGGACAAAGTAGAAAATATTAGCATGCATTGCATTTCATAAGTCATATGCACATCATG
GAATTTCATTTCCATTTTGTATGTGTATATGTGTGTAAACATATATACACATATGTAGACATACGTGTGTGTTTTGAATCATGATGTCA
AGTGTATTCATTACTGCAGACCACAGTCAAAGGGTTTTGAAAGCCACTGTTCCAATCCCTGCCAGCTCTCTGATTCTATAACTCTATTA
GATTACACTTGAGGAAGGTAAAATAATTCAATATATTTGATCATCCTCGCATATATAGACTTTTAGTTTAACGAGGAAAAAGTCTTGTA
TTGAAGAATAAAACTTGAAGAAAAATTTTAGCAGTGCTTTCAACCTTTAGAAATCTACAGTCAATATTTAGTTGTTTTTACCATTGTCA
GTATTTTCTATTCTGTGCTTTGATTTACTTCCATTCTAGTGTCTCTTGAGTAACATAACAGATTTATCTAAAATTCTTTATGCTGATAA

FIG.3A-23

CAAAGGCACTTCTATATAAAAACCTCCACATAAAATAAAATTATGGTTTTCAATTATACATTTTTATAACAATTATTACCACTTAAGAG
CATTTACTGGGTGTCAGGCAATGTTCTAAGACTTTTTCCATATATCAGATCATTTAATACCCTCAATGACCCTATAAGGGAAGTAGAAT
TCTTTCCCCAGTTTTTCAAATGAGGCACAGAGGAGGTTAAGCAACTTGTCTGAGCTCACACAGCTAGTAAATGGTAGAACTAGAATTCA
AACTCAAGCAGTATTTCTCTAGAATCAGTGAACGTAACCACTTTGCTAAACTGCCTGTGAAGTTACTTTTCTCAAAACAGCTCCTATTT
CACCATGTAAAGAAAAGTACAAACCCATAAAATAGCAAGTGCTGAAGAGAAGCCTTATGAAAGAAATATACAAATTCCAGCAAGTGAAA
ACGGTTGTGGTCCCTGGTTGTATAATAGTTACATGGGTGTTGACTTTACAATTATTTAAACCAAACATAAATACTTTATGCAGTTTTTA
TGTATGTTATACTCACAGAAAGAGAAGGGAAAAATTTTTAAATCATTCTCTTAAGGTTACATCAAGTTGCGTATCAGTTCAGTTCCATT
TAAATGATTCAAATCAAAGTCTGTGCATTTGAGAATTCATTAAGAGAGTAACATACATGTTATTCATTAAGAGTAACATAAATTTTGCA
TTGATTCTTGCCAAAATCACACCTACAACCATAAATTGTAAATTTCTAGGAAAACTCAGTACAAAACTTGGTGCAATGCAATAAAGTTT
GTGGCACAGACAGTAATACTCAGCAAACATCCCACCTCCTCTCTCATATTTTCCAGCTCCCCTTGTGGTTAAACGTTGCCATGTGGCAA
GTTCTGGCCAGTGAAGCGTGAGCAAAACTGAAAAGGGTTCTTTGTAGATTGAGACAGTGAAGAGCCTATGTGTGCTCATCTATTCTCTT
TTTCTGCTGAGGGCACAAAGAAAGTCCTGAAATCATGTGCTACAGCTATGAGATAATGTGCCTTTGCCTACCAGGCTTCTCAGTGTTTA
CTGGTGTGGAGCCCTTGTAATGGACACATAACATGAACAAGAAATAAATCTTTGTTGCATGAAGCCCTAGGAATGCCAGGACTAATCTG
TTACCTCAGCACAAACCCAGGCCTATCCTGACTAAGGTGGTATTAAATTACTATTGAATGTGTATTGGGATTTAGTAAACTTCTACTGT
ATAATCCTTCTTCTGTAGGTAGTTCCAAGGATTCATGAAGGAAATATTTCCAAACAAGATGAAACAATGATGACAGACTTAAGCATTCT
GCCTTCCTCTAATTTCACACTCAAGATCCCTCTTGAAGAAAGTGCTGAGAGTTCTAACTTCATTGGCTACGTAGTGGCAAAAGCTCTAC
AGCATTTTAAGGAACATTTTAAAACCTGGTAAGCAGAGTGCCTGGTTAGGAATGCCTTGTTGACAGGAATAGTTAATTCTCAAAAGGGA
AAAACAAAACTTGTTTCAAAATACCTGGAAAACATGTTTAACCTCATTAATAAAGACATGAAAACAAACAAGATGGCATTTTCTGCCTA
TCAGATTTGCAAATTAAAAAAAAACCCAGGAAATCCTGATAGGAATGTGATGAAATGGGAATTCTCATATATCATGTATTGGTGGGAAC
ATAATTGGTTTTGCATTTTTGAAAGCTATTTGATTATGCATATGAAGAGCCATAAAATTTCCTTTTGATATAATAATTCCACTTCCGAA
ATCAATCCTAAGGRATAAATCTAAATTTGATGAAMAKTCTCCCTCCAAGATCTAGATTTGCAGCATTATTTAAATATTAAAAGTTGGCC
GGGCGCAGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCGGGCGGATCACGAGGTCAGGAGATTGAGACCATCCTGGAT
AACACGGAGAAACTGCGTCTCTACTAAAATAAAAAAAATTAGCCGGGCATGGTGGCGGGCGCCTGTAGTCCCAGCTACTCGGGAGGCTG
AGGCAGGAGAATGGCGTGAACCCGGGAGGCAGARCTTGCAGTGAGCAGAGATCGCGCCACTGCACTCCAGCCTGGGCGACAGAGCAAGA
CTCTGTTTAAAAAAAAAAAAAAAAAAAAAAATATATATATATATATATATATATATATATATATATATATATGTTAAACATACTCTTAA
TGTGTAAAAACAAGAGAATGATTAAGTAKATTATGACTAAATACACTCAATACATTTTATGAAACGTTAAAAATATTCAAAAAATTTAA
ATAATGACTTGCTAACTACTTTAACAAGAGCTTTATTATCAGCTAGTCTTGGAGGTAATAGTATTATCATGATTTTTCAGAAAAAGATC
CTGAGGCTCAGTGTCCAAGGTCCAATGAACTACTCAGGTCGGAGGTGGTAGAGCAGCATGTGGAGCCAGTTCTCTCTCCGACTCCATCA
TCACACTGCACGGCTTCCTGTTAAGATATTTGCTCAAAAAATGCGAGATATAAAAATCTGGGTAATATGATCAACCTTAAAGAATAATT

FIG.3A–24

FIG.3A-25

GCACCACCACACTTGGCTAATTTTTGTATTTTTTAGTAGAGATGTGTTTTTACCATGTTGGCCGGGCTGGTCTCAAACTCCTGGCCTCA
AGTGATCCTTCCACCTCAGCCTCCCCAAGGGCTGGTATTACAGGCATGAGCCACTGCCTGGCCTGGCATTTCAATTTTTAAAATCTTCA
GTAATAAATGAAAATTTTTATCTTATTGTTATAATTTTTATGGTTTTTTATTATTCATGAGAATAAACATTTTCCAAGTTTGTTTACTG
ACTGAATTTCTTTTTTGTGCACCTTACTTGGTATCATGGATAAAATTTTGTCAATTTTCTGATTATATCAATGCATTCAGGGTCCCAAA
CCTGCCAAAGTTTAAAGAGAAAGATACTAAGGGAAAAACCAGGAAAAGATGGTAGAAAAGAATCACCCTGGCATTTTCAATCACGTAAA
CATTTGCTAGGTGCCCTAGCTGCAGGTATACAGCTCACTGAAACATGAATTCCAATTTTATAGGGTGAAATATATATTTAGAACCCTCT
TCTGGAACTTTCTTCTAGTTATCTAGCATCCTAAGTGCCTGGACGTTCCTGATTGGTTTGCAATGTGTTTTATTTCCCATCCCCAAGTT
TCATAGCTGCCGGCCCTGGGATCTACAGTCACAGGCTGTAACACAATATCTTGCACATCCTGAGTCTTTAATAAGCTTTTGTAGATGGG
CTCTTACCATCATCATCATCGTGAAAGGCAAATATACAAAATTTGTTGACTAATGTAATGAGTCATGAGTAACAGAAGTTTACTGACCA
AACACTACGTGCATGTAGAGTTCAGAATAAACACTTTATTATCACATCAGAGGAAAAGACCATCTTAGAGGCTCAACAACCCAGGAAAG
CTGTGACGATTTCTTCAAATTGTTAAGAATATCCATGCATATGGGTTTCACATTATTTTGCTACACACAGTACCAATTTTTCCAAAAGC
CAACAGCAGGTATTCTATTACCCATCCTGGACTTTTACTCCAAGAAAAAATACACTGAGTCTGTGAGTAATTTATTAGTATTTTGATCA
TTGCTGCTTTTTTTTTTTTTTTTAAGGTAAGAAGATCTAATGCATCCTATATCCAGTAAGTAGAATTATCTCTTCATCTGGGACCTGGAA
ATCCTGAAATAAAAAAGGATAATGCAATAAACACAGTTGCAGGAAAGTATGTTAGCTATATACTATGAAGTACTCTTAGTTTACTTATG
TTGAATGGCTTAGCTATTAATACTCAAATTGAGTTAAAATGAAAATTCCTCCTTAAAAAATCAAACGTAATATGTATTACATTTCATGG
TACATTAGTAGTTCTTTGTATATTGAATAAATACTAAATCACCTAGGTGTCTATGTTCTATCACATCTACAAACATGTCACTTCCTAAT
TAACAAAATGTTCTTCCTTTAGTTTGCTTTTGCACTTAAAATATATATAATTGACTTTTTTGGAAAAAAATCTAAGATTCATTGCTTTG
TTTTGTAAAGACCAATAGGTTCTGTATAGTCTTTTTTTAAATTGTGGTAAAATACACATGGCATTAATTTACCATTTTAACCATTTTAA
AGTGCACAATTTGTGGCATTAAGTACACTCACGTTGCTGTGCAACCATCACCACCGTCCATCTTCAGAACCTTTTTATCTTCCTAAACT
GAAACTCTGTACTCGTTAAGCACTCACTTCCCGTTTCCCCATCCCCCAGCCCGTAGCAACCACGACTGTACTTTCTATGAATTTGACTA
CTCTAGGTACTGCATGTAGGTGGAATCATACAGTATTTGTCTTTTGCTTCATTTTGTTTTGTTTTTTGTTTTCTAAGACAGGGTCTCAC
TCTGTCGCCCAGGCTGGAGTGCAGTGGTGCAATCACAGTGTCCTTTTGTGACTGGTTTATTTCACTTAGTGCCATGTTTTCAAGGTTCA
TCCATGTTGTTGCATGTCTCAGAACTTCCTTTTTTAGGCTAATATTCTTGCATGTATTTACCTAGTTTTGCTTATCCATTCAGCCATTG
ATGGACACTTGGGTTGCTTCCATCTTTTGGCTATTGTGAATAATGCTGTTTTGAACGTGGGTGTGCTACATAGTTACTTTTTAAAATTG
GCACAACAGCGCTGTCTTTTGACATACGTATTTTATGGAAAACACAAGATTTTCCTGGCTGACGCTCAACCTCATAATTTGGACCTTGG
TGCAACACAATAATAGGAGAGCTATGTGTCAGTATATATCACTAAGGATTACAATGAGAGTGTATACAGTCAGTATTACAAATTATAAA
AAGAAATGTAGGCCAGGCACGGTGCCTCACACCTGTAATCCCAGCACTTTGGGAGGCCAACGTGGGTGGATTACCTGAGGACAGGAGTT
CGAAACCAGCCTGGCCAACATGGTGAAAACCTGTATCTACTAAAAATACAAAAATTGGCCAGGTGTGGTGGCGCATGCCTGTAATCCCA
GCTACTCAGGAGGCTGAGATGGGAGAATTGCTTGAACCTGGGAGGCAGAGGTTGCACTGAGCCAAGATTGTGCCACTGTACTCCAGCCT

FIG.3A-26

GGGCAACAGAGCGAGACTCTTTTTTAATAAATAAATAAATAAATAAATATATAAAAGAAACGTAATGAAAGAGAGAGAACTCTGAACTT
TTAAAGAACTTTTCACCCAGTCTTGATCTATCTGACAGAAAGGCTTGTCAGAGAAAGTTAGAGTTCAGAGGCAGCCAATTGAATATAAT
TAACTCCAAATGAAGATAAACCTTTTCTAAATCATACTGAAGGCTATAAAAAATGAGAATTATGTTATTTTTTTTTTGAGACAGGGTCT
TACTCTATTGCCCAGGCTGGAGTGCAGTGGCATGATCTGGGCTCACTGAAGCCTGACCTCCTTGGCTCAGGTGATCCTCCCACCTCAGC
CTCCTGAGTAGCTGGGACTACAGGTACTACCATGCCCGTCTATTTTTGTATTTTTTTAGTAGAGATGGGGTTTCTCCATGTTGTCCAGG
CTGGTCTCAAACTCCCAGGCTCAAGCAATCTGCCCGCCTCAGCCTCCAAAAGTGCTGTAATTACAGGCATGAGCCACTGCTCCTGGCAG
GGAACTAATAGAATCCTGGGTTCTTCGGTGTGCAATAAAYCTCAAATACAGCTATTCAACCATAGATTTTAAATATTTGTTAGTGAAGG
TGACAAAAAAATAAGTGATTAAGAGAACCTATTTTCTATCCAATGAGCTATCAAAAGCTTATAGAGTGGAAAGAGAGTGGGGGAAGTGA
GGCTCAAAACAGCTAAATGGAAAGAAGATTTTGCATGCAGGCTGAACTGGATTTTCATCCTGGCTACTATATTCTCCAGATGTGTCACT
TTGGCCAAGATCCTTAATCTCAGTGTCATCTATAAGGTAATTAAAGTACACTAGTGCCCCACTAATCTGTGGTTTTGCTTTCCAAGCTT
TCAGTTACCCGAGATCAACTGCGGTTTTAAAATATTATGTGGAAAATTCCAGAAATACATAGTAAGTTTTCAATTGCATGCCATTAAAT
CTCATGCTGTCCCTGACCCCTTCCTCTCCGGAGGTGAATGCTCCCTTTGTCCAGTGGCTCCACGATGACTACATTCCCCAAATTGTTCT
CTTAGGAACCCTTTCTGTGTTCAAGGAACCCTTACTTTACTTAATTATGGCCCCAAAGCACAAGATAGGGATGCCGGCATACTGTTATA
ATTGTTCTATTTTATTATTAGTTATTGTTGTTCATCTCTACCTGTGACTAATTTATGAATTCAACTTTATCATAGGTATGTAGGTATAG
GAAAAAAACATGGTATGTATAAGGTTCAGTACTATCTGCAGTTTCAGACATCCCCTTGGGGTCTTGGAACATATCCCCCGTGGATAASG
GGAAACTACTGTAAAAGTTTGTSTTTTATAGAGTAGTTSTSAGAACTACATTAATCCATAATGTGTGSCTCATGATACTCATTGATAGA
TGGTAGTAGCAACAATAAAAAATAATATTATCAAGTAACTGATTCATAATTGACTCTCAAAAACGTTAATTTTCTGCTTTCCTTTACCT
AAGTTTACCTACATGTTTGAATTTGTAAAGGGAAGGTTTTTCTAGACCAATAATTTTCAAATATTTTTGCTCTCATACTTCCTCAAAGG
AAACTGAAAAAGTTGCAACATACTTGCATGTCATTTTTCTATATAAGTTGAAAGAATAGCAAATTGTTATTTTCCCACGCATCGTAAAG
ATTAGCAGGTCATCCCTCTTTAAAATGTACCAAATGGAATCTAAATATCATCGCAATTTGACCCAGCATCATCCATTTAAACAAATATA
CAAGTTTTTCTTTAACAATGAGAAATTTTATCTCATTACATTTTCTCCCTAAACTCTTATTTCAATCTACATTCCTAAGAATTTTATCC
TAATGTAGTATATTTTTATGCTTAAATATCTTTTGTTGATCAACACAATTTTGATCATTTTTAAATTTTAAAAATTAAGAACATCCTGT
GACATCAAATTCTAGGTATGAAATATTTATTCTAGATTGGGTGATCATTATAATTATTTTTTGTACATAATTGATCAAAATAACATAAA
TATACTACAAATTTCTATGACTACTAAACATATAAAAGTAAAATTTTAAACAAATATATCTCTTAATGAGAAGGAAGAGCTTTTTATAC
TCCAATAAGTTAACGTATCCACTAATAATTATTATTTCTTCCTAGAACAAGACAGGATTAAGCATCATGACCGTCCCTATTGGGGGATG
TTTTTATAGATGCAAGCACTGTGGCACCTACTGGTATAAATGCACCTGCTGATTGGAATGTTCTTTCCCCAGATCTTCCCCTGCTGGTT
TCTTCCCAGTATTCAGGTCTCAGCTCAAATGTGACTTCCTCAATGAGGCCTCCTGGTGATCAGATCTAAAGCACCCTCTACACAATCAC
TGTTTAGTGCTATACCCATTAATTTACTATCATCACACTTGTCACTATCTGCAGATGTCTTGTTTGGTTACTTTTGTNGTGTTTGTCAC
TGCCAGAATATCAGTTCTATGAAGAAAAGGGCCTTGTCTATTTTGACACTTATAGANATGATGNAGGNACGACATACAAATGGCCAATG

FIG.3A-27

GGCATATGGAAAAACGCTTGACTTCAAGAGTACTNATGGNTATNACCAACATTTATGGAGTAACTACTTTGAAAAGAACCATTCTGTCT
TTACTATCAAGCCAAGATACTCAAGGAAGGCAGCAGAAGTGGAAGCTCCATGTGGGCAGAGGAGCCTAGTCTTGAGATGTGATTTAGCT
GGTATTTGGGTGAAACAAATAAACCAGCCTCAAAATAACACAAGGGGCCGGGTGCAGTGGCTCACGCCTGTATCCCAGCACTTTGGGAG
GCTCGAGGCAGGCAGATTACTTCAGGTGAGGAGTTCGAGACCAGCCTGGCTAACATGGTGAACCTCCAT

FIG.3A-28

HKNG 1 mRNA expression in normal brain

Brain Regions	Gray Matter	White Matter	Neuron	Astrocytes	Oligodendrocytes
Frontal cortex(1)	+++	-	++	-	-
Motor cortex(2)	+++	-	++	-	-
Parietal cortex(3)	+++	-	++	-	-
Occipital cortex(4)	+++	-	++	-	-
Hippocampal formation(5)		-	++	-	-
CA1	+++	-	++	-	-
CA2	+++	-	++	-	-
CA3	na	na	na	na	na
CA4	+++	-	++	-	-
Dentate gyrus	++	-	+	-	-
subiculum	+++	-	++	-	-
parahippocampal gyri	+++	-	++	-	-
Caudate/Putamen(6)	+/-	-	+/-	-	-
GPi/GPe/Putamen(7)					
GPi	+	-	+	-	-
GPe	+	-	+	-	-
Putamen	+/-	-	+/-	-	-
Amygdala(8)	++	-	+	-	-
Thalamus(9)medial	++	-	+	-	-
Substantia nigra level(10)					
SNc(substantia nigra pars compacta)	++	-	++	-	-
SNr(substantia nigra pars reticulata)	+	-	+	-	-
Red Nucleus	+	-	+	-	-
3rd cranial nerve nuclei	+	-	+	-	-
superior colliculus	+	-	+	-	-
Upper pons(11)					
Locus coeruleus	+	-	+	-	-

FIG.4A

pontine nuclei	+++	-	++	-	-
Lower pons(12)					
locus coeruleus	+	-	+	-	-
pontine nuclei	+++	-	++	-	-
raphe nucleus(midline)	++	-	+	-	-
Medulla(13)					
Inferior olivary nucleus	++	-	+	-	-
12th cranial nerve nuclei	+	-	+	-	-
nucleus ambiguus(multipolar lower motor neurons)	+	-	+	-	-
Cerebellum(14)					
Purkinje cells	++	-	++	-	-
Granular layer	+	-	+/-	-	-
Molecular layer	+	-	+	-	-
Temporal pole(15)	+++	-	++	-	-
Cingulate cortex(16)	+++	-	++	-	-
Anterior thalamus(17)					
Subthalamic nucleus	?	-	?	-	-
Ventral anterior N.(VA),Ventral lateral N.(VL)	++	-	++	-	-
Hippocampal formation(18)					
CA1	na	na	na	na	na
CA2	na	na	na	na	na
CA3	+++	-	++	-	-
CA4	+++	-	++	-	-
subiculum	+++	-	++	-	-
parahippocampal gyri	+++	-	++	-	-
cervical cord (rostral position)					
anterior motor nuclei	++	-	+	-	-
sensory nuclei group	++	-	+	-	-

FIG.4B

pedigree	Affected Individuals	Phenotype	a.a. change	exon	comment	nt change	nt position
30124	3010189	scz	R331T	8	3 of 4 affected individuals	AGA -> ACT	51,641 51,642
	3010185	scz					
	3010184	scz					
30105	3010027	scz	I23T	3	the only affected individual	ATT -> ACT	35,044
31102	3110017	major depr	E202K	7	all three affected individual (also seen once in Costa Rica)	GAA -> AAA	45,487
	3110014	scz					
	3110003	scz					
30120	3010155	scz	E202K	7	one of the affected individuals	GAA -> AAA	45,487
30126	3010203	scz	intronic	10	3 of 4 affected individuals	insertion: GAATGCCTGGTTAG 21 base pairs 3' of exon 10	after 63,417
	3010210	scz					
	3010204	scz					
30140	3011486	scz	intronic	6	one of the two affected individuals	A -> T (24bp downstream of exon 6)	43,450
32301	3210041	scz			two of the three affected individuals		
	3210051	scz					

FIG.5A

pedigree	Affected Individuals	Phenotype	a.a. change	exon	comment	nt change	nt position
30120	3010155	scz	L34L	4	one of the two affected individuals	CTC -> CTA	36,307
32200	3210104	scz	L34L	4	both affected individuals	CTC -> CTA	36,307
	3210009	scz					
31109	3110013	scz	I23T	3	one of the two affected individuals	ATT -> ACT	35,044

FIG.5B

a.a. change	exon	nt change	position
non-coding 5'-UTR	1	G->C (35 bp upstream from 3' endo of exon 1)	15,385
L42L (silent)	4	CTG -> CTA	36,331
V123G	6	GTT -> GGT	43,184
non-coding (intronic)	6	A -> T (24 bp downstream from exon 6)	43,350
V301	7	GTC -> ATC	45,571

FIG.5C

AGTTGCGTCCCTCTCTGTTGCCAGGCTGGAGTTCAGTGGCATGTTCATAGCTC
ACTGAAGCCTCAAATTCNTGGGTTCAAGTGACCCTCCTACCTCAGCCCCATGA
GGACCTGGGACTACAGTTCCCTCCCTTTGGAACGCAGCGTGGGCACCTGCAA
CGCAGAGACCACTGTATCTCCGGTGCAGAATGTAATGAGTGCCTGATACATT
TGCCGAATAAACTATTCCAAGGGTTGAACTTGCTGGAAGCAANAGAAGCACT
ATTCTGGTAACAGCGGGAACATGAAGCCGCCACTCTTGGTGTTTATTGTGTGT
CTGCTGTGGTTGAAAGACAGTCACTGCGCACCCACTTGGAAGGACAAAACTG
CTATCAGTGAAAACCTGAAGAGTTTTTCTGA

FIG.6A

AGTTGCGTCCCTCTCTGTTGCCAGGCTGGAGTTCAGTGGCATGTTCTTAGCTC
ACTGAAGCCTCAAATTCCTGGGTTCAAGTGACCCTCCCACCTCAGCCCCATGA
GGACCTGGGACTACAGATGGAGTCTTGCTCTCGTTGCCCAGACTGGAGTGCA
CTGCTGCGATCTCAGCTCACTGCAACCTCTACCTCCCAGGTTCAAGCGATTCT
CCTGCCTCAGCCTCTCGAGTGGCTGGGACTATAGTAACAGCGGGAACATGAA
GCCGCCACTCTTGGTGTTTATTGTGTGTCCGCTGTGGTTGAAAGACAGTCACT
GCGCACCCACTTGGAAGGACAAAACTGCTATCAGTGAAAACCTGAAGAGTTT
TTCT

FIG.6B

```
CTTGGAGTCAACTGAGTGTGGACTGAAACTTCCAAAAACTGACATGAGGAGTCACTGGAGAATCATGATCAAGGAGCTA  79

                                                                    M   K   L      3
CACACTCTGACTTAACTTTATTCTGTGGACAATGAGAGACAACTGCAAGGATTAACAGTGAGAAC ATG AAG CTG    153

 P   L   L   M   F   P   V   C   L   L   W   L   K   D   C   H   C   A   P   T   23
CCA CTT TTG ATG TTT CCC GTG TGT CTG CTA TGG TTG AAA GAC TGT CAT TGT GCA CCT ACT  213

 W   K   D   K   T   A   I   S   E   N   A   N   S   F   S   E   A   G   E   I   43
TGG AAG GAC AAA ACT GCC ATC AGT GAA AAC GCG AAC AGT TTT TCT GAG GCT GGG GAG ATA  273

 D   V   D   G   E   V   K   I   A   L   I   G   I   K   Q   M   K   I   M   M   63
GAC GTA GAT GGA GAG GTG AAG ATA GCT TTG ATT GGC ATT AAA CAG ATG AAA ATC ATG ATG  333

 E   R   R   E   E   E   H   S   K   L   M   K   T   L   K   K   C   K   E   E   83
GAA AGG AGA GAG GAA GAA CAC AGC AAA CTA ATG AAA ACC TTG AAG AAG TGC AAA GAA GAA  393

 K   Q   E   A   L   K   L   M   N   E   V   H   E   H   L   E   E   E   E   S   103
AAG CAG GAG GCC CTG AAA CTT ATG AAT GAA GTT CAT GAA CAC CTG GAG GAG GAA GAA AGC  453

 L   C   Q   V   S   L   A   D   S   W   D   E   C   R   A   C   L   E   S   N   123
TTA TGC CAG GTT TCT CTG GCA GAT TCC TGG GAT GAA TGC AGG GCT TGC CTG GAA AGT AAC  513

 C   M   R   F   D   T   T   C   Q   P   A   W   S   S   V   K   N   M   V   E   143
TGC ATG AGG TTT GAT ACC ACC TGC CAA CCT GCA TGG TCC TCT GTG AAA AAT ATG GTG GAA  573

 Q   F   F   R   K   I   Y   Q   F   L   F   P   L   Q   E   N   D   R   S   G   163
CAG TTT TTC AGG AAG ATC TAT CAG TTT CTG TTT CCT CTC CAG GAA AAT GAC AGA AGT GGC  633

 P   V   S   K   G   V   T   E   E   D   A   Q   V   S   H   I   E   H   V   F   183
CCT GTC AGC AAA GGG GTC ACT GAG GAA GAT GCG CAG GTG TCA CAC ATA GAG CAT GTG TTC  693
```

FIG.7A

```
 S  Q  L  S  A  D  V  T  S  L  F  N  R  S  L  Y  V  F  K  Q   203
AGC CAG CTG AGC GCA GAT GTG ACA TCT CTC TTC AAC AGA AGC CTT TAC GTC TTC AAA CAG   753

 L  R  R  E  F  D  Q  A  F  Q  S  Y  F  T  S  G  T  D  V  T   223
CTG CGG CGA GAA TTT GAC CAG GCT TTT CAG TCA TAT TTC ACA TCG GGG ACT GAC GTT ACA   813

 E  P  F  F  F  P  S  L  S  K  E  P  A  Y  R  A  D  A  E  P   243
GAG CCT TTC TTT TTT CCA TCT TTG TCC AAG GAG CCA GCC TAC AGA GCA GAT GCT GAG CCA   873

 S  W  A  I  P  N  V  F  Q  L  L  C  N  L  S  F  S  V  Y  Q   263
AGC TGG GCC ATT CCC AAT GTC TTC CAG CTG CTC TGC AAC TTG AGT TTC TCA GTT TAT CAA   933

 S  V  S  E  K  L  I  T  T  L  R  A  T  E  D  P  P  K  Q  D   283
AGT GTC AGT GAA AAA CTC ATC ACA ACC CTG CGT GCC ACA GAG GAC CCT CCA AAA CAA GAC   993

 K  D  S  N  Q  G  G  P  I  S  K  I  L  P  E  Q  D  R  G  S   303
AAA GAC TCC AAC CAG GGA GGC CCG ATT TCA AAG ATA CTA CCT GAG CAA GAC AGA GGC TCA  1053

 D  G  K  L  G  Q  N  L  S  D  C  V  N  F  R  K  R  C  Q  K   323
GAT GGG AAA CTT GGC CAG AAT TTG TCT GAT TGC GTT AAT TTT CGC AAG AGA TGC CAG AAA  1113

 C  Q  D  Y  L  S  D  D  C  P  N  V  P  E  L  Y  R  E  L  N   343
TGC CAG GAT TAT CTA TCT GAT GCA TGC CCT AAT GTG CCT GAA CTA TAC AGA GAA CTC AAT  1173

 E  A  L  R  L  V  S  R  S  N  Q  Q  Y  D  Q  V  V  Q  M  T   363
GAG GCC CTC CGA CTG GTC AGT AGA TCC AAT CAG CAA TAC GAC CAG GTG GTG CAG ATG ACC  1233

 Q  Y  H  L  E  D  T  T  L  L  M  E  K  M  R  E  Q  F  G  W   383
CAG TAT CAC CTG GAA GAC ACC ACG CTT CTG ATG GAG AAG ATG AGA GAG CAG TTT GGC TGG  1293
```

FIG.7B

```
 V  S  E  L  A  Y  Q  S  P  G  A  E  D  I  F  N  P  V  K  V   403
GTT TCT GAA CTG GCA TAC CAG TCC CCA GGA GCT GAG GAC ATC TTT AAT CCA GTG AAA GTA  1353

 M  V  A  L  S  A  H  E  G  N  S  S  D  Q  D  D  T  V  V  P   423
ATG GTA GCC CTA AGT GCT CAT GAA GGA AAT TCT TCT GAT CAA GAT GAC ACA GTG GTT CCT  1413

 S  S  L  L  P  S  S  N  F  T  L  S  S  P  L  E  K  S  A  G   443
TCA AGC CTC CTG CCT TCC TCT AAC TTC ACA CTC AGC AGC CCT CTT GAA AAG AGT GCT GGC  1473

 N  A  N  F  I  D  H  V  V  E  K  V  L  Q  H  F  K  E  H  F   463
AAC GCT AAC TTC ATT GAT CAC GTG GTA GAG AAG GTT CTT CAG CAC TTT AAG GAG CAC TTT  1533

 K  T  W  *                                                     467
AAA ACT TGG TAA                                                1545

GAAGATTTAGTCCATCCTATAATCAGCAAGAATTACACCTTCGGCCAAGACCTGAGAATTCTGAAAATACAAAGCAGGC  1624

TAACACAATGAACACAGCTGCATGAAAGTTAGGTATATATTAGGAAGCACTATTGGTTTACTTTGTTGAATGGAAGTTT  1703

AATAGCTATTCAAATTGAGTTAATATAAAAATTTCTTCCTAAAAAGTAAAATGTACATATGTAGAATATGATGCATTAG  1782

TTCTTTGTATACTAAATAAATACTGAGTCCCCT                               1815
```

FIG.7C

```
CTTGGAGTCAACTGAGTGTGGACTGAAACTTCCAAAAACTGACATGAGGAGTCACTGGAGAATCATGATCAAGGAGCTA    79

                                                                    M   K   L        3
CACACTCTGACTTAACTTTATTCTGTGGACAATGAGAGACAACTGCAAGGATTAACAGTGAGAAC ATG AAG CTG      153

 P   L   L   M   F   P   V   C   L   L   W   L   K   D   C   H   C   A   P   T     23
CCA CTT TTG ATG TTT CCC GTG TGT CTG CTA TGG TTG AAA GAC TGT CAT TGT GCA CCT ACT     213

 W   K   D   K   T   A   I   S   E   N   A   N   S   F   S   E   A   G   E   I     43
TGG AAG GAC AAA ACT GCC ATC AGT GAA AAC GCG AAC AGT TTT TCT GAG GCT GGG GAG ATA     273

 D   V   D   G   E   V   K   I   A   L   I   G   I   K   Q   M   K   I   M   M     63
GAC GTA GAT GGA GAG GTG AAG ATA GCT TTG ATT GGC ATT AAA CAG ATG AAA ATC ATG ATG     333

 E   R   R   E   E   E   H   S   K   L   M   K   T   L   K   K   C   K   E   E     83
GAA AGG AGA GAG GAA GAA CAC AGC AAA CTA ATG AAA ACC TTG AAG AAG TGC AAA GAA GAA     393

 K   Q   E   A   L   K   L   M   N   E   V   H   E   H   L   E   E   E   E   S    103
AAG CAG GAG GCC CTG AAA CTT ATG AAT GAA GTT CAT GAA CAC CTG GAG GAG GAA GAA AGC    453

 L   C   Q   V   S   L   A   D   S   W   D   E   C   R   A   C   L   E   S   N    123
TTA TGC CAG GTT TCT CTG GCA GAT TCC TGG GAT GAA TGC AGG GCT TGC CTG GAA AGT AAC    513

 C   M   R   F   D   T   T   C   Q   P   A   W   S   S   V   K   N   M   E   N    143
TGC ATG AGG TTT GAT ACC ACC TGC CAA CCT GCA TGG TCC TCT GTG AAA AAT ATG GAA AAT    573

 D   R   S   G   P   V   S   K   G   V   T   E   E   D   A   Q   V   S   H   I    163
GAC AGA AGT GGC CCT GTC AGC AAA GGG GTC ACT GAG GAA GAT GCG CAG GTG TCA CAC ATA    633

 E   H   V   F   S   Q   L   S   A   D   V   T   S   L   F   N   R   S   L   Y    183
GAG CAT GTG TTC AGC CAG CTG AGC GCA GAT GTG ACA TCT CTC TTC AAC AGA AGC CTT TAC    693
```

FIG.8A

```
 V  F  K  Q  L  R  R  E  F  D  Q  A  F  Q  S  Y  F  T  S  G  203
GTC TTC AAA CAG CTG CGG CGA GAA TTT GAC CAG GCT TTT CAG TCA TAT TTC ACA TCG GGG  753

 T  D  V  T  E  P  F  F  F  P  S  L  S  K  E  P  A  Y  R  A  223
ACT GAC GTT ACA GAG CCT TTC TTT TTT CCA TCT TTG TCC AAG GAG CCA GCC TAC AGA GCA  813

 D  A  E  P  S  W  A  I  P  N  V  F  Q  L  L  C  N  L  S  F  243
GAT GCT GAG CCA AGC TGG GCC ATT CCC AAT GTC TTC CAG CTG CTC TGC AAC TTG AGT TTC  873

 S  V  Y  Q  S  V  S  E  K  L  I  T  T  L  R  A  T  E  D  P  263
TCA GTT TAT CAA AGT GTC AGT GAA AAA CTC ATC ACA ACC CTG CGT GCC ACA GAG GAC CCT  933

 P  K  Q  D  K  D  S  N  Q  G  G  P  I  S  K  I  L  P  E  Q  283
CCA AAA CAA GAC AAA GAC TCC AAC CAG GGA GGC CCG ATT TCA AAG ATA CTA CCT GAG CAA  993

 D  R  G  S  D  G  K  L  G  Q  N  L  S  D  C  V  N  F  R  K  303
GAC AGA GGC TCA GAT GGG AAA CTT GGC CAG AAT TTG TCT GAT TGC GTT AAT TTT CGC AAG 1053

 R  C  Q  K  C  Q  D  Y  L  S  D  D  C  P  N  V  P  E  L  Y  323
AGA TGC CAG AAA TGC CAG GAT TAT CTA TCT GAT GAC TGC CCT AAT GTG CCT GAA CTA TAC 1113

 R  E  L  N  E  A  L  R  L  V  S  R  S  N  Q  Q  Y  D  Q  V  343
AGA GAA CTC AAT GAG GCC CTC CGA CTG GTC AGT AGA TCC AAT CAG CAA TAC GAC CAG GTG 1173

 V  Q  M  T  Q  Y  H  L  E  D  T  T  L  L  M  E  K  M  R  E  363
GTG CAG ATG ACC CAG TAT CAC CTG GAA GAC ACC ACG CTT CTG ATG GAG AAG ATG AGA GAG 1233

 Q  F  G  W  V  S  E  L  A  Y  Q  S  P  G  A  E  D  I  F  N  383
CAG TTT GGC TGG GTT TCT GAA CTG GCA TAC CAG TCC CCA GGA GCT GAG GAC ATC TTT AAT 1293
```

FIG.8B

```
 P  V  K  V  M  V  A  L  S  A  H  E  G  N  S  S  D  Q  D  D  403
CCA GTG AAA GTA ATG GTA GCC CTA AGT GCT CAT GAA GGA AAT TCT TCT GAT CAA GAT GAC 1353

 T  V  V  P  S  S  L  L  P  S  S  N  F  T  L  S  S  P  L  E  423
ACA GTG GTT CCT TCA AGC CTC CTG CCT TCC TCT AAC TTC ACA CTC AGC AGC CCT CTT GAA 1413

 K  S  A  G  N  A  N  F  I  D  H  V  V  E  K  V  L  Q  H  F  443
AAG AGT GCT GGC AAC GCT AAC TTC ATT GAT CAC GTG GTA GAG AAG GTT CTT CAG CAC TTT 1473

 K  E  H  F  K  T  W  *                                        451
AAG GAG CAC TTT AAA ACT TGG TAA                                1497

GAAGATTTAGTCCATCCTATAATCAGCAAGAATTACACCTTCGGCCAAGACCTGAGAATTCTGAAAATACAAAGCAGGC 1576

TAACACAATGAACACAGCTGCATGAAAGTTAGGTATATATTAGGAAGCACTATTGGTTTACTTTGTTGAATGGAAGTTT 1655

AATAGCTATTCAAATTGAGTTAATATAAAAATTTCTTCCTAAAAAGTAAAATGTACATATGTAGAATATGATGCATTAG 1734

TTCTTTGTATACTAAATAAATACTGAGTCCCCT                              1767
```

FIG.8C

```
CTTGGAGTCAACTGAGTGTGGACTGAAACTTCCAAAAACTGACATGAGGAGTCACTGGAGAATCATGATCAAGGAGCTA   79

                                                                    M   K   L       3
CACACTCTGACTTAACTTTATTCTGTGGACAATGAGAGACAACTGCAAGGATTAACAGTGAGAAC ATG AAG CTG     153

 P   L   L   M   F   P   V   C   L   L   W   L   K   D   C   H   C   A   P   T    23
CCA CTT TTG ATG TTT CCC GTG TGT CTG CTA TGG TTG AAA GAC TGT CAT TGT GCA CCT ACT   213

 W   K   D   K   T   A   I   S   E   N   A   N   S   F   S   E   A   G   E   I    43
TGG AAG GAC AAA ACT GCC ATC AGT GAA AAC GCG AAC AGT TTT TCT GAG GCT GGG GAG ATA   273

 D   V   D   G   E   V   K   I   A   L   I   G   I   K   Q   M   K   I   M   M    63
GAC GTA GAT GGA GAG GTG AAG ATA GCT TTG ATT GGC ATT AAA CAG ATG AAA ATC ATG ATG   333

 E   R   R   E   E   E   H   S   K   L   M   K   T   L   K   K   C   K   E   E    83
GAA AGG AGA GAG GAA GAA CAC AGC AAA CTA ATG AAA ACC TTG AAG AAG TGC AAA GAA GAA   393

 K   Q   E   A   L   K   L   M   N   E   V   H   E   H   L   E   E   E   E   S   103
AAG CAG GAG GCC CTG AAA CTT ATG AAT GAA GTT CAT GAA CAC CTG GAG GAG GAA GAA AGC   453

 L   C   Q   V   S   L   A   D   S   W   D   E   C   R   A   C   L   E   S   N   123
TTA TGC CAG GTT TCT CTG GCA GAT TCC TGG GAT GAA TGC AGG GCT TGC CTG GAA AGT AAC   513

 C   M   R   F   D   T   T   C   Q   P   A   W   S   S   V   K   N   M   E   P   143
TGC ATG AGG TTT GAT ACC ACC TGC CAA CCT GCA TGG TCC TCT GTG AAA AAT ATG GAG CCA   573

 A   Y   R   A   D   A   E   P   S   W   A   I   P   N   V   F   Q   L   L   C   163
GCC TAC AGA GCA GAT GCT GAG CCA AGC TGG GCC ATT CCC AAT GTC TTC CAG CTG CTC TGC   633

 N   L   S   F   S   V   Y   Q   S   V   S   E   K   L   I   T   T   L   R   A   183
AAC TTG AGT TTC TCA GTT TAT CAA AGT GTC AGT GAA AAA CTC ATC ACA ACC CTG CGT GCC   693
```

FIG.9A

```
 T   E   D   P   P   K   Q   D   K   D   S   N   Q   G   G   P   I   S   K   I  203
ACA GAG GAC CCT CCA AAA CAA GAC AAA GAC TCC AAC CAG GGA GGC CCG ATT TCA AAG ATA 753

 L   P   E   Q   D   R   G   S   D   G   K   L   G   Q   N   L   S   D   C   V  223
CTA CCT GAG CAA GAC AGA GGC TCA GAT GGG AAA CTT GGC CAG AAT TTG TCT GAT TGC GTT  813

 N   F   R   K   R   C   Q   K   C   Q   D   Y   L   S   D   D   C   P   N   V  243
AAT TTT CGC AAG AGA TGC CAG AAA TGC CAG GAT TAT CTA TCT GAT GAC TGC CCT AAT GTG  873

 P   E   L   Y   R   E   L   N   E   A   L   R   L   V   S   R   S   N   Q   Q  263
CCT GAA CTA TAC AGA GAA CTC AAT GAG GCC CTC CGA CTG GTC AGT AGA TCC AAT CAG CAA  933

 Y   D   Q   V   V   Q   M   T   Q   Y   H   L   E   D   T   T   L   L   M   E  283
TAC GAC CAG GTG GTG CAG ATG ACC CAG TAT CAC CTG GAA GAC ACC ACG CTT CTG ATG GAG  993

 K   M   R   E   Q   F   G   W   V   S   E   L   A   Y   Q   S   P   G   A   E  303
AAG ATG AGA GAG CAG TTT GGC TGG GTT TCT GAA CTG GCA TAC CAG TCC CCA GGA GCT GAG 1053

 D   I   F   N   P   V   K   V   M   V   A   L   S   A   H   E   G   N   S   S  323
GAC ATC TTT AAT CCA GTG AAA GTA ATG GTA GCC CTA AGT GCT CAT GAA GGA AAT TCT TCT 1113

 D   Q   D   D   T   V   V   P   S   S   L   L   P   S   S   N   F   T   L   S  343
GAT CAA GAT GAC ACA GTG GTT CCT TCA AGC CTC CTG CCT TCC TCT AAC TTC ACA CTC AGC 1173

 S   P   L   E   K   S   A   G   N   A   N   F   I   D   H   V   V   E   K   V  363
AGC CCT CTT GAA AAG AGT GCT GGC AAC GCT AAC TTC ATT GAT CAC GTG GTA GAG AAG GTT 1233

 L   Q   H   F   K   E   H   F   K   T   W   *                                  375
CTT CAG CAC TTT AAG GAG CAC TTT AAA ACT TGG TAA                                1269
```

FIG.9B

GAAGATTTAGTCCATCCTATAATCAGCAAGAATTACACCTTCGGCCAAGACCTGAGAATTCTGAAAATACAAAGCAGGC 1348

TAACACAATGAACACAGCTGCATGAAAGTTAGGTATATATTAGGAAGCACTATTGGTTTACTTTGTTGAATGGAAGTTT 1427

AATAGCTATTCAAATTGAGTTAATATAAAAATTTCTTCCTAAAAAGTAAAATGTACATATGTAGAATATGATGCATTAG 1506

TTCTTTGTATACTAAATAAATACTGAGTCCCCT 1539

FIG.9C

```
CTTGGAGTCAACTGAGTGTGGACTGAAACTTCCAAAAACTGACATGAGGAGTCACTGGAGAATCATGATCAAGGAGCTA 79

                                                                  M  K  L      3
CACACTCTGACTTAACTTTATTCTGTGGACAATGAGAGACAACTGCAAGGATTAACAGTGAGAAC ATG AAG CTG    153

 P  L  L  M  F  P  V  C  L  L  W  L  K  D  C  H  C  A  P  T     23
CCA CTT TTG ATG TTT CCC GTG TGT CTG CTA TGG TTG AAA GAC TGT CAT TGT GCA CCT ACT   213

 W  K  D  K  T  A  I  S  E  N  A  N  S  F  S  E  A  G  E  I     43
TGG AAG GAC AAA ACT GCC ATC AGT GAA AAC GCG AAC AGT TTT TCT GAG GCT GGG GAG ATA   273

 D  V  D  G  E  V  K  I  A  L  I  G  I  K  Q  M  K  I  M  M     63
GAC GTA GAT GGA GAG GTG AAG ATA GCT TTG ATT GGC ATT AAA CAG ATG AAA ATC ATG ATG   333

 E  R  R  E  E  E  H  S  K  L  M  K  T  L  K  K  C  K  E  E     83
GAA AGG AGA GAG GAA GAA CAC AGC AAA CTA ATG AAA ACC TTG AAG AAG TGC AAA GAA GAA   393

 K  Q  E  A  L  K  L  M  N  E  V  H  E  H  L  E  E  E  E  S    103
AAG CAG GAG GCC CTG AAA CTT ATG AAT GAA GTT CAT GAA CAC CTG GAG GAG GAA GAA AGC   453

 L  C  Q  V  S  L  A  D  S  W  D  E  C  R  A  C  L  E  S  N    123
TTA TGC CAG GTT TCT CTG GCA GAT TCC TGG GAT GAA TGC AGG GCT TGC CTG GAA AGT AAC   513

 C  M  R  F  D  T  T  C  Q  P  A  W  S  S  V  K  N  M  P  A    143
TGC ATG AGG TTT GAT ACC ACC TGC CAA CCT GCA TGG TCC TCT GTG AAA AAT ATG CCA GCC   573

 Y  R  A  D  A  E  P  S  W  A  I  P  N  V  F  Q  L  L  C  N    163
TAC AGA GCA GAT GCT GAG CCA AGC TGG GCC ATT CCC AAT GTC TTC CAG CTG CTC TGC AAC   633

 L  S  F  S  V  Y  Q  S  V  S  E  K  L  I  T  T  L  R  A  T    183
TTG AGT TTC TCA GTT TAT CAA AGT GTC AGT GAA AAA CTC ATC ACA ACC CTG CGT GCC ACA   693
```

FIG.10A

```
 E  D  P  P  K  Q  D  K  D  S  N  Q  G  G  P  I  S  K  I  L  203
GAG GAC CCT CCA AAA CAA GAC AAA GAC TCC AAC CAG GGA GGC CCG ATT TCA AAG ATA CTA  753

 P  E  Q  D  R  G  S  D  G  K  L  G  Q  N  L  S  D  G  V  N  223
CCT GAG CAA GAC AGA GGC TCA GAT GGG AAA CTT GGC CAG AAA TTG TCT GAT TGC GTT AAT  813

 F  R  K  R  C  Q  K  C  Q  D  Y  L  S  D  D  C  P  N  V  P  243
TTT CGC AAG AGA TGC CAG AAA TGC CAG GAT TAT CTA TCT GAT GAC TGC CCT AAT GTG CCT  873

 E  L  Y  R  E  L  N  E  A  L  R  L  V  S  R  S  N  Q  Q  Y  263
GAA CTA TAC AGA GAA CTC AAT GAG GCC CTC CGA CTG GTC AGT AGA TCC AAT CAG CAA TAC  933

 D  Q  V  V  Q  M  T  Q  Y  H  L  E  D  T  T  L  L  M  E  K  283
GAC CAG GTG GTG CAG ATG ACC CAG TAT CAC CTG GAA GAC ACC ACG CTT CTG ATG GAG AAG  993

 M  R  E  Q  F  G  W  V  S  E  L  A  Y  Q  S  P  G  A  E  D  303
ATG AGA GAG CAG TTT GGC TGG GTT TCT GAA CTG GCA TAC CAG TCC CCA GGA GCT GAG GAC 1053

 I  F  N  P  V  K  V  M  V  A  L  S  A  H  E  G  N  S  S  D  323
ATC TTT AAT CCA GTG AAA GTA ATG GTA GCC CTA AGT GCT CAT GAA GGA AAT TCT TCT GAT 1113

 Q  D  D  T  V  V  P  S  S  L  L  P  S  S  N  F  T  L  S  S  343
CAA GAT GAC ACA GTG GTT CCT TCA AGC CTC CTG CCT TCC TCT AAC TTC ACA CTC AGC AGC 1173

 P  L  E  K  S  A  G  N  A  N  F  I  D  H  V  V  E  K  V  L  363
CCT CTT GAA AAG AGT GCT GGC AAC GCT AAC TTC ATT GAT CAC GTG GTA GAG AAG GTT CTT 1233

 Q  H  F  K  E  H  F  K  T  W  *                              374
CAG CAC TTT AAG GAG CAC TTT AAA ACT TGG TAA                   1266
```

FIG.10B

GAAGATTTAGTCCATCCTATAATCAGCAAGAATTACACCTTCGGCCAAGACCTGAGAATTCTGAAAATACAAAGCAGGC 1345

TAACACAATGAACACAGCTGCATGAAAGTTAGGTATATATTAGGAAGCACTATTGGTTTACTTTGTTGAATGGAAGTTT 1424

AATAGCTATTCAAATTGAGTTAATATAAAAATTTCTTCCTAAAAAGTAAAATGTACATATGTAGAATATGATGCATTAG 1503

TTCTTTGTATACTAAATAAATACTGAGTCCCCT 1536

FIG.10C

```
                                                        M  K  P  P  L  L  V  F      8
GCAACCTCGTTGGTGAGAGCCTGCAGTTAGTGTCACGGCGGAAAC ATG AAG CCG CCA CTC TTG GTG TTT    69

 I  V  Y  L  L  R  L  R  D  C  Q  C  A  P  T  G  K  D  R  T   28
ATT GTG TAT CTG CTG CGG CTG AGA GAC TCT CAG TGT GCG CCT ACA GGG AAG GAC CGA ACT  129

 S  I  R  E  D  P  K  G  F  S  K  A  G  E  I  D  V  D  E  E   48
TCC ATC CGT GAA GAC CCG AAG GGT TTT TCC AAG GCT GGG GAG ATA GAC GTA GAT GAA GAG  189

 V  K  K  A  L  I  G  M  K  Q  M  K  I  L  M  E  R  R  E  E   68
GTG AAG AAG GCT TTG ATT GGC ATG AAG CAG ATG AAA ATC CTG ATG GAA AGA AGA GAG GAG  249

 E  H  S  K  L  M  R  T  L  K  K  C  R  E  E  K  Q  E  A  L   88
GAA CAT AGC AAA CTA ATG AGA ACA CTG AAG AAA TGC AGA GAA GAA AAG CAG GAG GCC CTG  309

 K  L  M  N  E  V  Q  E  H  L  E  E  E  E  R  L  C  Q  V  S  108
AAG CTT ATG AAT GAA GTT CAA GAA CAT CTA GAA GAG GAA GAA AGG CTA TGC CAG GTG TCT  369

 L  M  G  S  W  D  E  C  K  S  C  L  E  S  D  C  M  R  F  Y  128
CTG ATG GGT TCC TGG GAC GAA TGC AAA TCT TGC CTG GAA AGT GAC TGC ATG AGA TTT TAT  429

 T  T  C  Q  S  S  W  S  S  M  K  S  T  I  E  R  V  F  R  K  148
ACA ACC TGC CAA AGC AGT TGG TCC TCT ATG AAA TCC ACG ATT GAA CGG GTT TTC CGG AAG  489

 I  Y  Q  F  L  F  P  F  H  E  D  D  E  K  E  L  P  V  G  E  168
ATA TAT CAG TTT CTC TTT CCT TTC CAT GAA GAC GAT GAA AAA GAG CTT CCT GTT GGT GAG  549

 K  F  T  E  E  D  V  Q  L  M  Q  I  E  N  V  F  S  Q  L  T  188
AAG TTC ACT GAG GAA GAT GTA CAG CTG ATG CAG ATA GAG AAT GTG TTC AGC CAG CTG ACC  609
```

FIG.11A

```
 V   D   V   G   F   L   Y   N   M   S   F   H   V   F   K   Q   M   Q   Q   E  208
GTG GAT GTG GGA TTT CTC TAT AAC ATG AGC TTT CAC GTC TTC AAA CAG ATG CAG CAA GAA  669

 F   D   L   A   F   Q   S   Y   F   M   S   D   T   D   S   M   E   P   Y   F  228
TTT GAC CTG GCT TTT CAA TCA TAC TTT ATG TCA GAC ACA GAC TCC ATG GAG CCT TAC TTT  729

 F   P   A   F   S   K   E   P   A   K   K   A   H   P   M   Q   S   W   D   I  248
TTT CCA GCT TTT TCC AAA GAG CCA GCA AAA AAA GCA CAT CCT ATG CAG AGT TGG GAC ATT  789

 P   S   F   F   Q   L   F   C   N   F   S   L   S   V   Y   Q   S   V   S   A  268
CCC AGC TTC TTC CAG CTG TTT TGT AAT TTC AGC CTC TCT GTT TAT CAA AGT GTC AGC GCA  849

 T   V   T   E   M   L   K   A   I   E   D   L   S   K   Q   D   K   D   S   A  288
ACA GTT ACA GAG ATG CTG AAG GCC ATT GAG GAC TTA TCC AAA CAA GAC AAA GAT TCT GCC  909

 H   G   G   P   S   S   T   T   W   P   V   R   G   R   G   L   C   G   E   P  308
CAC GGT GGA CCG AGT TCC ACG ACG TGG CCT GTG CGG GGC AGA GGG CTG TGT GGA GAA CCT  969

 G   Q   N   S   S   E   C   L   Q   F   H   A   R   C   Q   K   C   Q   D   Y  328
GGC CAG AAC TCG TCC GAA TGT CTC CAA TTT CAT GCA AGA TGC CAG AAA TGT CAG GAT TAC 1029

 L   W   A   D   C   P   A   V   P   E   L   Y   T   K   A   D   E   A   L   E  348
CTA TGG GCA GAC TGC CCT GCT GTT CCT GAA CTA TAC ACA AAG GCG GAT GAG GCC CTT GAG 1089

 L   V   N   I   S   N   Q   Q   Y   A   Q   V   L   Q   M   T   Q   H   H   L  368
TTG GTC AAC ATA TCC AAT CAG CAG TAT GCC CAG GTA CTC CAG ATG ACC CAG CAT CAC TTG 1149

 E   D   T   T   Y   L   M   E   K   M   R   E   Q   F   G   W   V   T   E   L  388
GAG GAC ACC ACG TAT CTG ATG GAG AAG ATG AGA GAG CAG TTT GGT TGG GTA ACA GAG CTG 1209
```

FIG.11B

```
 A   S   Q   T   P   G   S   E   N   I   F   S   F   I   K   V   V   P   G   V  408
GCC AGC CAG ACC CCA GGA AGC GAG AAC ATC TTC AGT TTC ATA AAG GTA GTT CCA GGT GTT 1269

 H   E   G   N   F   S   K   Q   D   E   K   M   I   D   I   S   I   L   P   S  428
CAC GAA GGA AAT TTC TCC AAA CAA GAT GAA AAG ATG ATA GAC ATA AGC ATT CTG CCT TCC 1329

 S   N   F   T   L   T   I   P   L   E   E   S   A   E   S   S   D   F   I   S  448
TCT AAT TTC ACA CTC ACC ATC CCT CTT GAA GAA AGT GCT GAG AGT TCC GAC TTC ATT AGC 1289

 Y   M   L   A   K   A   V   Q   H   F   K   E   H   F   K   S   W   *          466
TAC ATG CTG GCC AAA GCT GTA CAG CAT TTT AAG GAA CAT TTT AAA TCT TGG TAA         1443

GCAGAGTATTTGATTAGGGACGTTTGCTGATAGGAATAGATGGTTCTTAAAAGGGAAAAATGACAAAACTAGCTTTTGA 1522
ATACCTTGAAAACGTATTCAACCTCATTAATAATCAAAGGCATGAAAACTAAGACAAGTTAGCAGTTTTTACCTATTGA 1601
ATTTTCAAATTAAAAAAAAAAATCCTGATAGAATGCAATGAAATGAGAATTCTTATATGTGATTGCCAGAAACAAACTG 1680
GTTTTGTCTTTTTGAAAAGTTATTCAATTATACATATCAAGAGTCATCAAATTTCTTTTTAATATAATAATTCCACTTC 1759
TGGAATCAATCCAAAGGAGTAAATCTAAAATTGAATTGAAGTTCCCACCCCAAGATCAATATTTGCAAATTATTTAAAA 1838
TAGTAAACTGTTAAAAACTGAATGTCATCTGAATGTCTAAAAACCAGAAATGGTTAAAAGCTGTGGCTAAATATGCTCC 1917
AAATATCTTATAAAACCATTAAAAATATTTATAAAATTTAAATCATGACATGACATCTGCTGGAACAAGAGTTTATTCT 1996
AAGCCTATCTATAAGGCAAATATTATTATTACTATCTTCCAGAAAAGAAACTTGAGACTCAGGGTCCAAGTGTTAGTTG 2075
CTCAGTCATGTCTGACTCTTTGGGACCCCTTGGACTGTAGCCCACCAGGCTCCTCTGTCCGTGGGATTCTTCAGACAGG 2154
AATACTGGGGCAGGTTGCTATTTCCTTCTCCAGGAAATCTTCCCTATCCAGGGATGGAACCCAGGTCTCCTGCATTGCA 2233
GGTAGATGCTTTACTATCTGAGCAACCAAATGAATTACTCAAGTCAGTAGGGGGTAGAGGCAAATTTTAACTTAGTTTT 2312
CTCTGAATCATAATTGCCACATTAAACTGGTTCCTGTTGGGACATTTGGTTGAAAAAAATAAAGTGAAAAATGAGTATA 2391
AAACTCTATAAATGTAATGATCAAAACGAAAAAAAATCTACAATCTGCATTAAAAATAAAAAGGGTTGGCAGG       2464
```

FIG.11C

```
                                                                M   K   P   P   L      5
CAGAAGCTGGTGGCAACCTCGTTGGTGAGAGCCTGCAGTTAGTGTCACGGCGGAAAC ATG AAG CCG CCA CTC     72

 L   V   F   I   V   Y   L   L   R   L   R   D   C   Q   C   A   P   T   G   K    25
TTG GTG TTT ATT GTG TAT ATG ATG CGG CTG AGA GAC TGT CAG TGT GCG CCT ACA GGG AAG  132

 D   R   T   S   I   R   E   D   P   K   G   F   S   K   A   G   E   I   D   V    45
GAC CGA ACT TCC ATC CGT GAA GAC CCG AAG GGT TTT TCC AAG GCT GGG GAG ATA GAC GTA  192

 D   E   E   V   K   K   A   L   I   G   M   K   Q   M   K   I   L   M   E   R    65
GAT GAA GAG GTG AAG AAG GCT TTG ATT GGC ATG AAG CAG ATG AAA ATC CTG ATG GAA AGA  252

 R   E   E   E   H   S   K   L   M   R   T   L   K   K   C   R   E   E   K   Q    85
AGA GAG GAG GAA CAT AGC AAA CTA ATG GAG ACA CTG AAG AAA TGC AGA GAA GAA AAG CAG  312

 E   A   L   K   L   M   N   E   V   Q   E   H   L   E   E   E   E   R   L   C   105
GAG GCC CTG AAG CTT ATG AAT GAA GTT CAA GAA CAT CTA GAA GAG GAA GAA AGG CTA TGC  372

 Q   V   S   L   M   G   S   W   D   E   C   K   S   C   L   E   S   D   C   M   125
CAG GTG TCT CTG ATG GGT TCC TGG GAC GAA TGC AAA TCT TGC CTG GAA AGT GAC TGC ATG  432

 R   F   Y   T   T   C   Q   S   S   W   S   S   M   K   S   T   I   E   R   V   145
AGA TTT TAT ACA ACC TGC CAA AGC AGT TGG TCC TCT ATG AAA TCC ACG ATT GAA CGG GTT  492

 F   R   K   I   Y   Q   F   L   F   P   F   H   E   D   D   E   K   E   L   P   165
TTC CGG AAG ATA TAT CAG TTT CTC TTT CCT TTC CAT GAA GAC GAT GAA AAA GAG CTT CCT  552

 V   G   E   K   F   T   E   E   D   V   Q   L   M   Q   I   E   N   V   F   S   185
GTT GGT GAG AAG TTC ACT GAG GAA GAT GTA CAG CTG ATG CAG ATA GAG AAT GTG TTC AGC  612
```

FIG.12A

```
 Q  L  T  V  D  V  G  F  L  Y  N  M  S  F  H  V  F  K  Q  M  205
CAG CTG ACC GTG GAT GTG GGA TTT CTC TAT AAC ATG AGC TTT CAC GTC TTC AAA CAG ATG  672

 Q  Q  E  F  D  L  A  F  Q  S  Y  F  M  S  D  T  D  S  M  E  225
CAG CAA GAA TTT GAC CTG GCT TTT CAA TCA TAC TTT ATG TCA GAC ACA GAC TCC ATG GAG  732

 P  Y  F  F  P  A  F  S  K  E  P  A  K  K  A  H  P  M  Q  S  245
CCT TAC TTT TTT CCA GCT TTT TCC AAA GAG CCA GCA AAA AAA GCA CAT CCT ATG CAG AGT  792

 W  D  I  P  S  F  F  Q  L  F  C  N  F  S  L  S  V  Y  Q  S  265
TGG GAC ATT CCC AGC TTC TTC CAG CTG TTT TGT AAT TTC AGC CTC TCT GTT TAT CAA AGT  852

 V  S  A  T  V  T  E  M  L  K  A  I  E  D  L  S  K  Q  D  K  285
GTC AGC GCA ACA GTT ACA GAG ATG CTG AAG GCC ATT GAG GAC TTA TCC AAA CAA GAC AAA  912

 D  S  A  H  G  G  P  S  S  T  T  W  P  V  R  G  R  G  L  C  305
GAT TCT GCC CAC GGT GGA CCG AGT TCC ACG ACG TGG CCT GTG CGG GGC AGA GGG CTG TGT  972

 G  E  P  G  Q  N  S  S  E  C  L  Q  F  H  A  R  C  Q  K  C  325
GGA GAA CCT GGC CAG AAC TCG TCC GAA TGT CTC CAA TTT CAT GCA AGA TGC CAG AAA TGT 1032

 Q  D  Y  L  W  A  D  C  P  A  V  P  E  L  Y  T  K  A  D  E  345
CAG GAT TAC CTA TGG GCA GAC TGC CCT GCT GTT CCT GAA CTA TAC ACA AAG GCG GAT GAG 1092

 A  L  E  L  V  N  I  S  N  Q  Q  Y  A  Q  V  L  Q  M  T  Q  365
GCC CTT GAG TTG GTC AAC ATA TCC AAT CAG CAG TAT GCC CAG GTA CTC CAG ATG ACC CAG 1152

 H  H  L  E  D  T  T  Y  L  M  E  K  M  R  E  Q  F  G  W  V  385
CAT CAC TTG GAG GAC ACC ACG TAT CTG ATG GAG AAG ATG AGA GAG CAG TTT GGT TGG GTA 1212
```

FIG.12B

```
 T   E   L   A   S   Q   T   P   G   S   E   N   I   F   S   F   I   K   V   V   405
ACA GAG CTG GCC AGC CAG ACC CCA GGA AGC GAG AAC ATC TTC AGT TTC ATA AAG GTA GTT 1272

 P   G   V   H   E   G   N   F   S   K   Q   D   E   K   M   I   D   I   S   I   425
CCA GGT GTT CAC GAA GGA AAT TTC TCC AAA CAA GAT GAA AAG ATG ATA GAC ATA AGC ATT 1332

 L   P   S   S   N   F   T   L   T   I   P   L   E   E   S   A   E   S   S   D   445
CTG CCT TCC TCT AAT TTC ACA CTC ACC ATC CCT CTT GAA GAA AGT GCT GAG AGT TCC GAC 1392

 F   I   S   Y   M   L   A   K   A   V   Q   H   F   K   E   H   F   K   S   W   465
TTC ATT AGC TAC ATG CTG GCC AAA GCT GTA CAG CAT TTT AAG GAA CAT TTT AAA TCT TGG 1452

 *                                                                               466
TAA                                                                              1455

GCAGAGTATTTGATTAGGGACGTTTGCTGATAGGAATAGATGGTTCTTAAAAGGGAAAAATGACAAAACTAGCTTTTGA 1534

ATACCTTGAAAACGTATTCAACCTCATTAATAATCAAAGGCATGAAAACTAAGACAAGTTAGCAGTTTTTACCTATTGA 1613

ATTTTCAAATTAAAAAAAAAAATCCTGATAGAATGCAATGAAATGAGAATTCTTATATGTGATTGCCAGAAACAAACTG 1692

GTTTTGTCTTTTTGAAAAGTTATTCAATTATACATATCAAGAGTCATCAAATTTCTTTTTAATATAATAATTCCACTTC 1771

TGGAATCAATCCAAAGGAGTAAATCTAAAATTGAATTGAAGTTCCCACCCCAAGATCAATATTTGCAAATTATTTAAAA 1850

TAGTAAACTGTTAAAAACTGAATGTCATCTGAATGTCTAAAAACCAGAAATGGTTAAAAGCTGTGGCTAAATATGCTCC 1929

AAATATCTTATAAAACCATTAAAAATATTTATAAAATTTAAATCATGACATGACATCTGCTGGAACAAGAGTTTATTCT 2008

AAGCCTATCTATAAGGCAAATATTATTATTACTATCTTCCAGAAAAGAAACTTGAGACTCAGGGTCCAAGTGTTAGTTG 2087

CTCAGTCATGTCTGACTCTTTGAGACCCCTTGGACTGTGGCCCACCAGGCTCCTCTGTCCATGGGATTCTTCAGACAAG 2166
```

FIG.12C

AATACTGGAGCAGGTTGCTATTTCCTTCTCCAGGAAATCTTCCCTATCCAGGGATGGAACCCAGGTCTCCTGCATTGCA 2245

GGTAGATGCTTTACTATCTGAGCAACCAAATGAATTACTCAAGTCAGTAGGGGGTAGAGGCAAATTTTAACTTAGTTTT 2324

CTCTGAATCATAATTGCCACATTAAACTGGTTCCTGTTGGGACATTTGGTTGAAAAAAATAAAGTGAAAAATGAGTATA 2403

AAACTCTATAAATGTAATGATCAAAACGAAAAAAAATCTACAATCTGCATTAAAAATAAAAAGGGTTGGCAGGAATTAC 2482

GGTTGGAAATGGATGATTTTTTTTAACCTTTTCATCTTTTGATATTTTACAATTTTCTATAATGAATAAATAATTTTGA 2561

GATTTCAAATTAGAAGATATGTTGCTAAAATAGCTAGGTAAATGTAGATTGAACACTGTATCAATGTGTTCTCATCTTT 2640

AAACTTTAGTATAAGTACTTCTATTCCATGGTAATCCTACAGTAAGACGAAATGTAAATCTGTTCGGTCTACAGGAAAA 2719

ACAACTAAATGACATTTCAGACGTACATTACCATCTCTGTTAGGATAATCTTCTGAATTAATGGCACAATTAGAACTGT 2798

ACATAGTATTCTCCTTTGGTAAAATGGTCAATCTTAAAGAAGCATTAAATGTTAATTCTAAGTTATTACTCATAAGGGA 2877

CCTTGTAGGTAGGTCCCTATCAATGTATAATTAAGCTGGGTATTTCTAGATTCGCTGCCTCTCCCTTTATCTCTGAATG 2956

TTGGAGAGGTTGTTGGTCATCAATCAACCAATATCTTTTTAGCATCTTCTAAGTGAAGGC 3016

FIG.12D

```
                                                                  M   K      2
GTGAAGGTCCTTACAGAAGCTGGTGGCAACCTCGTTGGTGAGAGCCTGCAGTTAGTGTCACGGCGGAAAC ATG AAG     76

 P   P   I   L   V   F   I   V   Y   L   L   Q   L   R   D   C   Q   C   A   P    22
CCG CCA ATC TTG GTG TTT ATC GTG TAT CTG CTG CAG CTG AGA GAC TGT CAG TGT GCG CCT  136

 T   G   K   D   R   T   S   I   R   E   D   P   K   G   F   S   K   A   G   E    42
ACA GGG AAG GAC CGA ACT TCC ATC CGT GAA GAC CCG AAG GGT TTT TCC AAG GCT GGG GAG  196

 I   D   V   D   E   E   V   K   K   A   L   I   G   M   K   Q   M   K   I   L    62
ATA GAC GTA GAT GAA GAG GTG AAG AAG GCT TTG ATT GGC CTG AAG CAG ATG AAA ATC CTG  256

 M   E   R   R   E   E   E   H   S   K   L   M   R   T   L   K   K   C   R   E    82
ATG GAA AGA AGA GAG GAG GAA CAT AGC AAA CTA ATG AGA ACC CTG AAG AAA TGC AGA GAA  316

 E   K   Q   E   A   L   K   L   M   N   E   V   Q   E   H   L   E   E   E   E   102
GAA AAG CAG GAG GCC CTG AAG CTT ATG AAT GAA GTT CAA GAA CAT CTA GAA GAG GAA GAA  376

 R   L   C   Q   V   S   L   M   G   S   W   D   E   C   K   S   C   L   E   S   122
AGG CTA TGC CAG GTG TCT CTG ATG GGT TCC TGG GAC GAA TGC AAA TCT TGC CTG GAA AGT  436

 D   C   M   R   F   Y   T   T   C   Q   S   S   W   S   S   M   K   S   T   I   142
GAC TGC ATG AGA TTT TAT ACA ACC TGC CAA AGC AGT TGG TCC TCT ATG AAA TCC ACG ATT  496

 E   R   V   F   R   K   I   Y   Q   F   L   F   P   F   H   E   D   D   E   K   162
GAA CGG GTT TTC CGG AAG ATA TAT ACG TTT CTC TTT CCT TTC CAT GAA GAC GAT GAA AAA  556

 E   L   P   V   G   E   K   F   T   E   E   D   V   Q   L   M   Q   I   E   N   182
GAG CTT CCT GTT GGT GAG AAG TTC ACT GAG GAA GAT GTA CAG CTG ATG CAG ATA GAG AAT  616
```

FIG.13A

```
 V   F   S   Q   L   T   V   D   V   G   F   L   Y   N   M   S   F   H   V   F  202
GTG TTC AGC CAG CTG ACC GTG GAC GTG GGA TTT CTC TAT AAC ATG TGC TTT CAC GTC TTC  676

 K   Q   M   Q   Q   E   F   D   L   A   F   Q   S   Y   F   M   S   D   T   D  222
AAA CAG ATG CAG CAA GAA TTT GAC CTG GCT TTT CAA TCA TAC TTT ATG TCA GAC ACA GAC  736

 S   M   E   P   Y   F   F   P   A   F   S   K   E   P   A   K   K   A   H   P  242
TCC ATG GAG CCT TAC TTT TTT CCA GCT TTT TCC AAA GAG CCA GCA AAA AAA GCA CAT CCT  796

 M   Q   S   W   D   I   P   S   F   F   Q   L   F   C   N   F   S   L   S   V  262
ATG CAG AGT TGG GAC ATT CCC AGC TTC TTC CAG CTG TTT TGT AAT TTC AGC CTC TCT GTT  856

 Y   Q   S   V   S   A   T   V   T   E   M   L   K   A   I   E   D   L   S   K  282
TAT CAA AGT GTC AGC GCA ACA GTT ACA GAG ATG CTG AAG GCC ATT GAG GAC TTA TCC AAA  916

 Q   D   K   D   S   A   H   G   G   P   S   S   T   T   W   P   V   R   G   R  302
CAA GAC AAA GAT TCT GCC CAC GGT GGA CCG AGT TCC ACG ACG TGG CCT GTG CGG GGC AGA  976

 G   L   C   G   E   P   G   Q   N   S   S   E   C   L   Q   F   H   A   R   C  322
GGG CTG TGT GGA GAA CTT GGC CAG AAC TCG TCC GAA TGT CTC CAA TTT CAT GCA AGA TGC 1036

 Q   K   C   Q   D   Y   L   W   A   D   C   P   A   V   P   E   L   Y   T   K  342
CAG AAA TGT CAG GAT TAC CTA TGG GCA GAC TGC CCT GCT GTT CCT GAA CTA TAC ACA AAG 1096

 A   D   E   A   L   E   L   V   N   I   S   N   Q   Q   Y   A   Q   V   L   Q  362
GCG GAT GAG GCC CTT GAG TTG GTC AAC ATA TCC CCT CAG CAG TAT GCC CAG GTA CTC CAG 1156

 M   T   Q   H   H   L   E   D   T   T   Y   L   M   E   K   M   R   E   Q   F  382
ATG ACC CAG CAT CAC TTG GAG GAC ACC ACG TAT CTG ATG GAG AAG ATG AGA GAG CAG TTT 1216
```

FIG.13B

```
 G   W   V   T   E   L   A   S   Q   T   P   G   S   E   N   I   F   S   F   I   402
GGT TGG GTA ACA GAG CTG GCC AGC CAG ACC CCA GGA AGC GAG AAC ATC TTC AGT TTC ATA 1276

 K   V   V   P   G   V   H   E   G   N   F   S   K   Q   D   E   K   M   I   D   422
AAG GTA GTT CCA GGT GTT CAC GAA GGA AAT TTC TCC AAA CAA GAT GAA AAG ATG ATA GAC 1336

 I   S   I   L   P   S   S   N   F   T   L   T   I   P   L   E   E   S   A   E   442
ATA AGC ATT CTG CCT TCC TCT AAT TTC ACA CTC ACC ATC CCT CTT GAA GAA AGT GCA GAG 1396

 S   S   D   F   I   S   Y   M   L   A   K   A   V   Q   H   F   K   E   H   F   462
AGT TCC GAC TTC ATT AGC TAC ATG CTG GCC AAA GCT GTA CAG CAT TTT AAG GAA CAT TTT 1456

 K   S   W   *                                                                   466
AAA TCT TGG TAA                                                                 1468

GCAGAGTATTTGATTAGGGACGTTTGCTGATAGGAATAGATGGTTCTTAAAAGGGAAAAATGACAAAACTAGCTTTTGA 1547
ATACCTTGAAAACGTATTCAACCTCATTAATAATCAAAGGCATGAAAACTAAGACAAGTTAGCAGTTTTTACCTATTGA 1626
ATTTTCAAATTAAAAAAAAAATCCTGATAGAATGCAATGAAATGAGAATTCTTATATGTGATTGCCAGAAACAAACTGG 1705
TTTTGTCTTTTTGAAAAGTTATTCAATTATACATATCAAGAGTCATCAAATTTCTTTTTAATATAATAATTCCACTTCT 1784
GGAATCAATCCAAAGGAGTAAATCTAAAATTGAATTGAAGTTCCCACCCCAAGATCAATATTTGCAAATTATTTAAAAT 1863
AGTAAACTGTTAAAAACTGAATGTCATCTGAATGTCTAAAAACCAGAAATGGTTAAAAGCTGTGGCTAAATATGCTCCA 1942
AATATCTTATAAAACCATTAAAAATATTTATAAAATTTAAATCATGACATGACATCTGCTGGAACAAGAGTTTATTCTA 2021
AGCCTATCTATAAGGCAAATATTATTATTACTATCTTCCAGAAAAGAAACTTGAGACTCAGGGTCCAAGTGTTAGTTGC 2100
TCAGTCATGTCTGACTCTTTGAGACCCCTTGGACTGTAGCCCACCAGGCTCCTCTGTCCATGGGATTCTTCAGACAAGA 2179
ATACTGGAGCAGGTTGCTATTTCCTTCTCCAGGAAATCTTCCCTATCCAGGGATGGAACCCAGGTCTCCTGCATTGCAG 2258
GTAGATGCTTTACTATCTGAGCAACCAAATGAATTACTCAAGTCAGTAGGGGGTAGAGGCAAATTTTAACTTAGTTTTC 2337
TCTGAATCATAATTGCCACATTAAACTGGTTCCTGTTGGGACATTTGGTTGAAAAAAATAAAGTGAAAAATGAGTATAA 2416
AACTCTATAAATGTAATGATCAAAACGAAAAAAAATCTACAATCTGCATTAAAAATAAAAAGGGTTGGCAGG         2488
```

FIG.13C

```
      CTTGGAGTCAACTGAGTGTGGACTGAAACTTCCAAAAACT  Majority
                10        20        30        40
1     CTTGGAGTCAACTGAGTGTGGACTGAAACTTCCAAAAACT  gphkng1815-1.
1     CTTGGAGTCAACTGAGTGTGGACTGAAACTTCCAAAAACT  gp7b-1.
1     CTTGGAGTCAACTGAGTGTGGACTGAAACTTCCAAAAACT  gp7c-1.
1     CTTGGAGTCAACTGAGTGTGGACTGAAACTTCCAAAAACT  gp7d-1.

      GACATGAGGAGTCACTGGAGAATCATGATCAAGGAGCTAC  Majority
                50        60        70        80
41    GACATGAGGAGTCACTGGAGAATCATGATCAAGGAGCTAC  gphkng1815-1.
41    GACATGAGGAGTCACTGGAGAATCATGATCAAGGAGCTAC  gp7b-1.
41    GACATGAGGAGTCACTGGAGAATCATGATCAAGGAGCTAC  gp7c-1.
41    GACATGAGGAGTCACTGGAGAATCATGATCAAGGAGCTAC  gp7d-1.

      ACACTCTGACTTAACTTTATTCTGTGGACAATGAGAGACA  Majority
                90        100       110       120
81    ACACTCTGACTTAACTTTATTCTGTGGACAATGAGAGACA  gphkng1815-1.
81    ACACTCTGACTTAACTTTATTCTGTGGACAATGAGAGACA  gp7b-1.
81    ACACTCTGACTTAACTTTATTCTGTGGACAATGAGAGACA  gp7c-1.
81    ACACTCTGACTTAACTTTATTCTGTGGACAATGAGAGACA  gp7d-1.

      ACTGCAAGGATTAACAGTGAGAACATGAAGCTGCCACTTT  Majority
                130       140       150       160
121   ACTGCAAGGATTAACAGTGAGAACATGAAGCTGCCACTTT  gphkng1815-1.
121   ACTGCAAGGATTAACAGTGAGAACATGAAGCTGCCACTTT  gp7b-1.
121   ACTGCAAGGATTAACAGTGAGAACATGAAGCTGCCACTTT  gp7c-1.
121   ACTGCAAGGATTAACAGTGAGAACATGAAGCTGCCACTTT  gp7d-1.
```

FIG.14A

```
    TGATGTTTCCCGTGTGTCTGCTATGGTTGAAAGACTGTCA  Majority
                170       180       190       200
161 TGATGTTTCCCGTGTGTCTGCTATGGTTGAAAGACTGTCA  gphkng1815-1.
161 TGATGTTTCCCGTGTGTCTGCTATGGTTGAAAGACTGTCA  gp7b-1.
161 TGATGTTTCCCGTGTGTCTGCTATGGTTGAAAGACTGTCA  gp7c-1.
161 TGATGTTTCCCGTGTGTCTGCTATGGTTGAAAGACTGTCA  gp7d-1.

    TTGTGCACCTACTTGGAAGGACAAAACTGCCATCAGTGAA  Majority
                210       220       230       240
201 TTGTGCACCTACTTGGAAGGACAAAACTGCCATCAGTGAA  gphkng1815-1.
201 TTGTGCACCTACTTGGAAGGACAAAACTGCCATCAGTGAA  gp7b-1.
201 TTGTGCACCTACTTGGAAGGACAAAACTGCCATCAGTGAA  gp7c-1.
201 TTGTGCACCTACTTGGAAGGACAAAACTGCCATCAGTGAA  gp7d-1.

    AACGCGAACAGTTTTTCTGAGGCTGGGGAGATAGACGTAG  Majority
                250       260       270       280
241 AACGCGAACAGTTTTTCTGAGGCTGGGGAGATAGACGTAG  gphkng1815-1.
241 AACGCGAACAGTTTTTCTGAGGCTGGGGAGATAGACGTAG  gp7b-1.
241 AACGCGAACAGTTTTTCTGAGGCTGGGGAGATAGACGTAG  gp7c-1.
241 AACGCGAACAGTTTTTCTGAGGCTGGGGAGATAGACGTAG  gp7d-1.

    ATGGAGAGGTGAAGATAGCTTTGATTGGCATTAAACAGAT  Majority
                290       300       310       320
281 ATGGAGAGGTGAAGATAGCTTTGATTGGCATTAAACAGAT  gphkng1815-1.
281 ATGGAGAGGTGAAGATAGCTTTGATTGGCATTAAACAGAT  gp7b-1.
281 ATGGAGAGGTGAAGATAGCTTTGATTGGCATTAAACAGAT  gp7c-1.
281 ATGGAGAGGTGAAGATAGCTTTGATTGGCATTAAACAGAT  gp7d-1.
```

FIG.14B

```
      GAAAATCATGATGGAAAGGAGAGAGGAAGAACACAGCAAA  Majority
                330       340       350       360
321   GAAAATCATGATGGAAAGGAGAGAGGAAGAACACAGCAAA  gphkng1815-1.
321   GAAAATCATGATGGAAAGGAGAGAGGAAGAACACAGCAAA  gp7b-1.
321   GAAAATCATGATGGAAAGGAGAGAGGAAGAACACAGCAAA  gp7c-1.
321   GAAAATCATGATGGAAAGGAGAGAGGAAGAACACAGCAAA  gp7d-1.

      CTAATGAAAACCTTGAAGAAGTGCAAAGAAGAAAAGCAGG  Majority
                370       380       390       400
361   CTAATGAAAACCTTGAAGAAGTGCAAAGAAGAAAAGCAGG  gphkng1815-1.
361   CTAATGAAAACCTTGAAGAAGTGCAAAGAAGAAAAGCAGG  gp7b-1.
361   CTAATGAAAACCTTGAAGAAGTGCAAAGAAGAAAAGCAGG  gp7c-1.
361   CTAATGAAAACCTTGAAGAAGTGCAAAGAAGAAAAGCAGG  gp7d-1.

      AGGCCCTGAAACTTATGAATGAAGTTCATGAACACCTGGA  Majority
                410       420       430       440
401   AGGCCCTGAAACTTATGAATGAAGTTCATGAACACCTGGA  gphkng1815-1.
401   AGGCCCTGAAACTTATGAATGAAGTTCATGAACACCTGGA  gp7b-1.
401   AGGCCCTGAAACTTATGAATGAAGTTCATGAACACCTGGA  gp7c-1.
401   AGGCCCTGAAACTTATGAATGAAGTTCATGAACACCTGGA  gp7d-1.

      GGAGGAAGAAAGCTTATGCCAGGTTTCTCTGGCAGATTCC  Majority
                450       460       470       480
441   GGAGGAAGAAAGCTTATGCCAGGTTTCTCTGGCAGATTCC  gphkng1815-1.
441   GGAGGAAGAAAGCTTATGCCAGGTTTCTCTGGCAGATTCC  gp7b-1.
441   GGAGGAAGAAAGCTTATGCCAGGTTTCTCTGGCAGATTCC  gp7c-1.
441   GGAGGAAGAAAGCTTATGCCAGGTTTCTCTGGCAGATTCC  gp7d-1.
```

FIG.14C

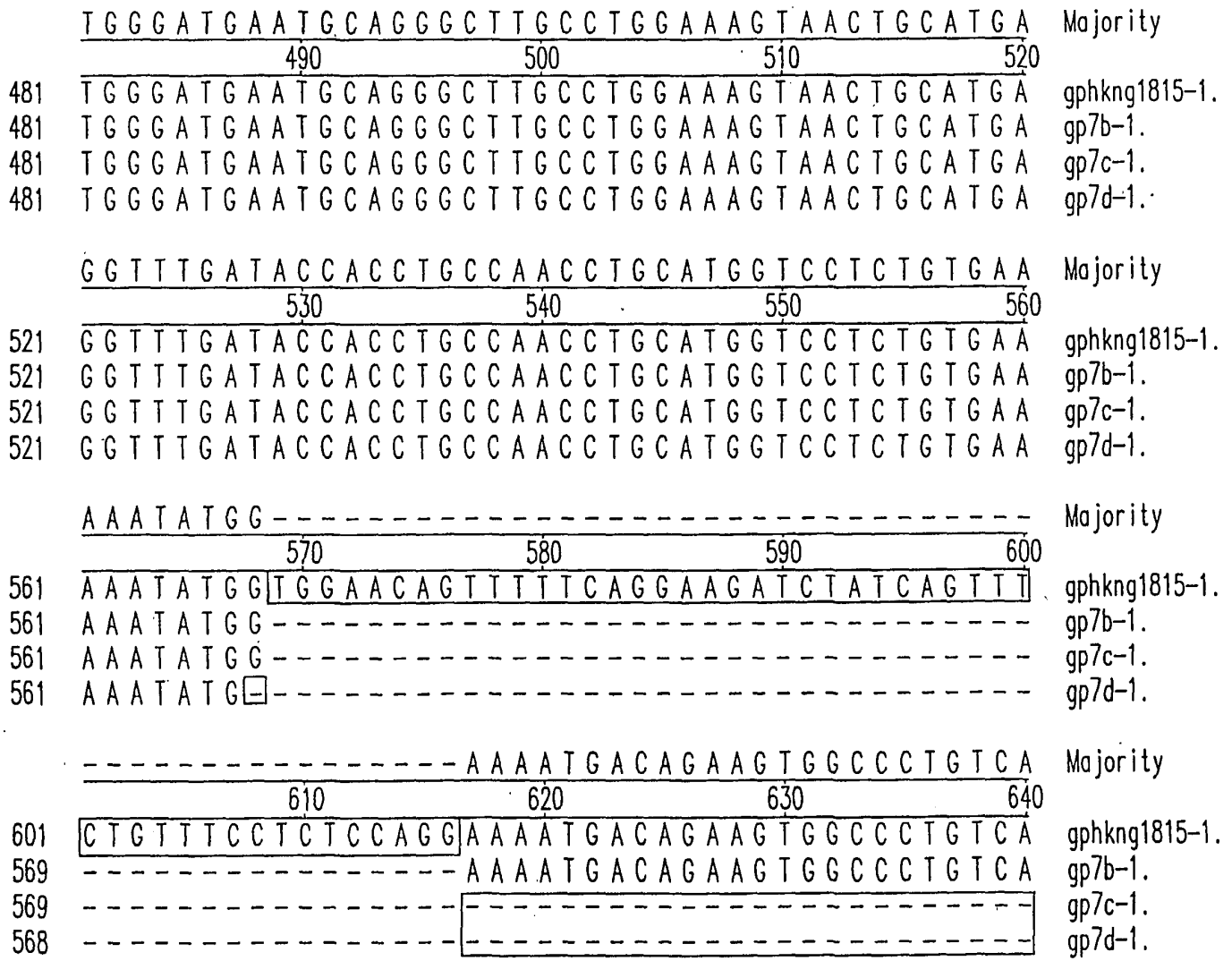

TGGGATGAATGCAGGGCTTGCCTGGAAAGTAACTGCATGA Majority
490 500 510 520
481 TGGGATGAATGCAGGGCTTGCCTGGAAAGTAACTGCATGA gphkng1815-1.
481 TGGGATGAATGCAGGGCTTGCCTGGAAAGTAACTGCATGA gp7b-1.
481 TGGGATGAATGCAGGGCTTGCCTGGAAAGTAACTGCATGA gp7c-1.
481 TGGGATGAATGCAGGGCTTGCCTGGAAAGTAACTGCATGA gp7d-1.
GGTTTGATACCACCTGCCAACCTGCATGGTCCTCTGTGAA Majority
530 540 550 560
521 GGTTTGATACCACCTGCCAACCTGCATGGTCCTCTGTGAA gphkng1815-1.
521 GGTTTGATACCACCTGCCAACCTGCATGGTCCTCTGTGAA gp7b-1.
521 GGTTTGATACCACCTGCCAACCTGCATGGTCCTCTGTGAA gp7c-1.
521 GGTTTGATACCACCTGCCAACCTGCATGGTCCTCTGTGAA gp7d-1.
AAATATGG Majority
570 580 590 600
561 AAATATGGTGGAACAGTTTTTCAGGAAGATCTATCAGTTT gphkng1815-1.
561 AAATATGG gp7b-1.
561 AAATATGG gp7c-1.
561 AAATATG gp7d-1.
AAAATGACAGAAGTGGCCCTGTCA Majority
610 620 630 640
601 CTGTTTCCTCTCCAGGAAAATGACAGAAGTGGCCCTGTCA gphkng1815-1.
569 AAAATGACAGAAGTGGCCCTGTCA gp7b-1.
569 gp7c-1.
568 gp7d-1.

FIG.14D

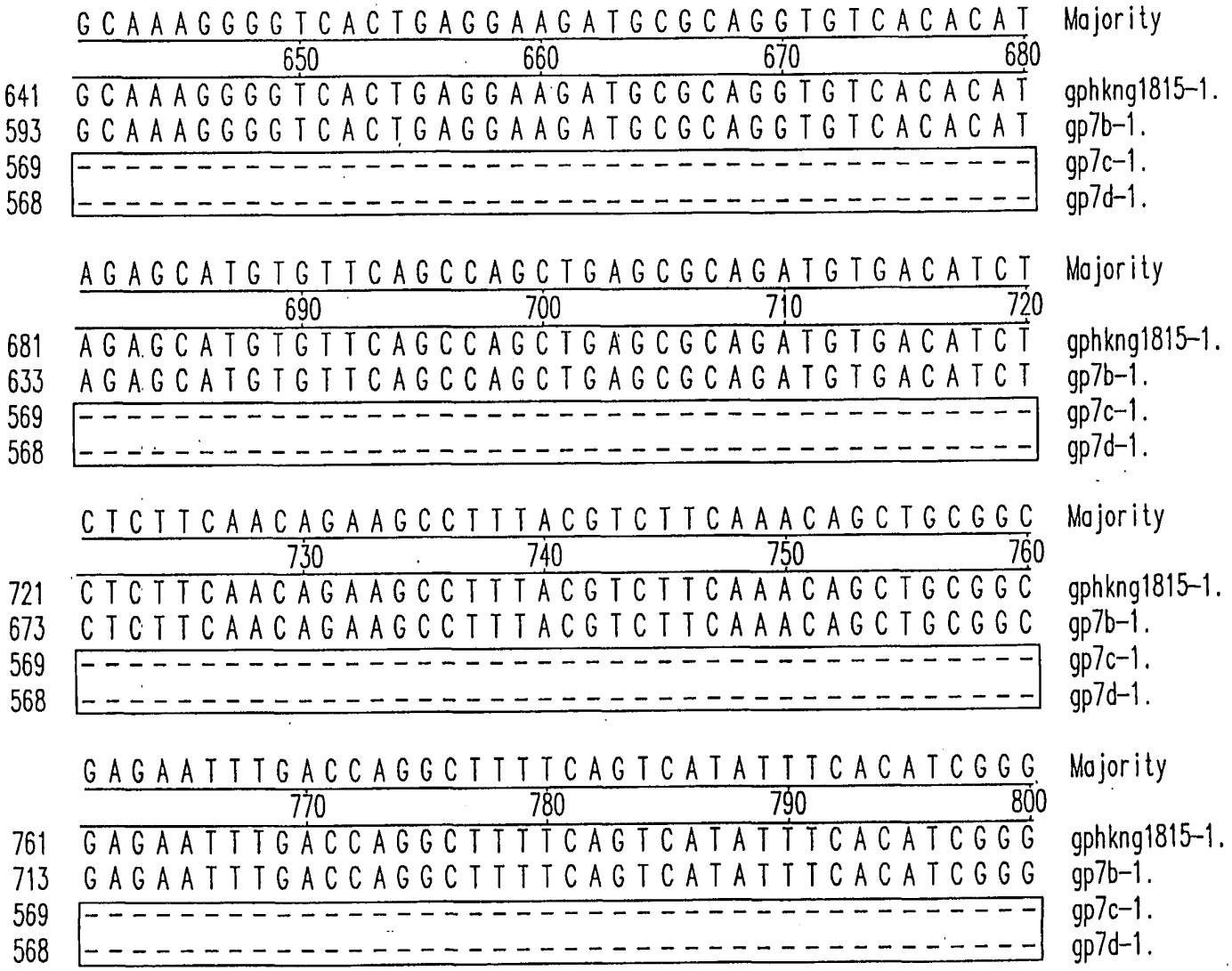

GCAAAGGGGTCACTGAGGAAGATGCGCAGGTGTCACACAT Majority
650 660 670 680
641 GCAAAGGGGTCACTGAGGAAGATGCGCAGGTGTCACACAT gphkng1815-1.
593 GCAAAGGGGTCACTGAGGAAGATGCGCAGGTGTCACACAT gp7b-1.
569 gp7c-1.
568 gp7d-1.
AGAGCATGTGTTCAGCCAGCTGAGCGCAGATGTGACATCT Majority
690 700 710 720
681 AGAGCATGTGTTCAGCCAGCTGAGCGCAGATGTGACATCT gphkng1815-1.
633 AGAGCATGTGTTCAGCCAGCTGAGCGCAGATGTGACATCT gp7b-1.
569 gp7c-1.
568 gp7d-1.
CTCTTCAACAGAAGCCTTTACGTCTTCAAACAGCTGCGGC Majority
730 740 750 760
721 CTCTTCAACAGAAGCCTTTACGTCTTCAAACAGCTGCGGC gphkng1815-1.
673 CTCTTCAACAGAAGCCTTTACGTCTTCAAACAGCTGCGGC gp7b-1.
569 gp7c-1.
568 gp7d-1.
GAGAATTTGACCAGGCTTTTCAGTCATATTTCACATCGGG Majority
770 780 790 800
761 GAGAATTTGACCAGGCTTTTCAGTCATATTTCACATCGGG gphkng1815-1.
713 GAGAATTTGACCAGGCTTTTCAGTCATATTTCACATCGGG gp7b-1.
569 gp7c-1.
568 gp7d-1.

FIG.14E

```
      GACTGACGTTACAGAGCCTTTCTTTTTTCCATCTTTGTCC  Majority
                 810       820       830       840
801   GACTGACGTTACAGAGCCTTTCTTTTTTCCATCTTTGTCC  gpkng1815-1.
753   GACTGACGTTACAGAGCCTTTCTTTTTTCCATCTTTGTCC  gp7b-1.
569   ----------------------------------------  gp7c-1.
568   ----------------------------------------  gp7d-1.

      AAGGAGCCAGCCTACAGAGCAGATGCTGAGCCAAGCTGGG  Majority
                 850       860       870       880
841   AAGGAGCCAGCCTACAGAGCAGATGCTGAGCCAAGCTGGG  gpkng1815-1.
793   AAGGAGCCAGCCTACAGAGCAGATGCTGAGCCAAGCTGGG  gp7b-1.
569   ----AGCCAGCCTACAGAGCAGATGCTGAGCCAAGCTGGG  gp7c-1.
568   ------CCAGCCTACAGAGCAGATGCTGAGCCAAGCTGGG  gp7d-1.

      CCATTCCCAATGTCTTCCAGCTGCTCTGCAACTTGAGTTT  Majority
                 890       900       910       920
881   CCATTCCCAATGTCTTCCAGCTGCTCTGCAACTTGAGTTT  gpkng1815-1.
833   CCATTCCCAATGTCTTCCAGCTGCTCTGCAACTTGAGTTT  gp7b-1.
605   CCATTCCCAATGTCTTCCAGCTGCTCTGCAACTTGAGTTT  gp7c-1.
602   CCATTCCCAATGTCTTCCAGCTGCTCTGCAACTTGAGTTT  gp7d-1.

      CTCAGTTTATCAAAGTGTCAGTGAAAAACTCATCACAACC  Majority
                 930       940       950       960
921   CTCAGTTTATCAAAGTGTCAGTGAAAAACTCATCACAACC  gpkng1815-1.
873   CTCAGTTTATCAAAGTGTCAGTGAAAAACTCATCACAACC  gp7b-1.
645   CTCAGTTTATCAAAGTGTCAGTGAAAAACTCATCACAACC  gp7c-1.
642   CTCAGTTTATCAAAGTGTCAGTGAAAAACTCATCACAACC  gp7d-1.
```

FIG.14F

```
      CTGCGTGCCACAGAGGACCCTCCAAAACAAGACAAAGACT  Majority
                970       980       990       1000
961   CTGCGTGCCACAGAGGACCCTCCAAAACAAGACAAAGACT  gphkng1815-1.
913   CTGCGTGCCACAGAGGACCCTCCAAAACAAGACAAAGACT  gp7b-1.
685   CTGCGTGCCACAGAGGACCCTCCAAAACAAGACAAAGACT  gp7c-1.
682   CTGCGTGCCACAGAGGACCCTCCAAAACAAGACAAAGACT  gp7d-1.

      CCAACCAGGGAGGCCCGATTTCAAAGATACTACCTGAGCA  Majority
                1010      1020      1030      1040
1001  CCAACCAGGGAGGCCCGATTTCAAAGATACTACCTGAGCA  gphkng1815-1.
953   CCAACCAGGGAGGCCCGATTTCAAAGATACTACCTGAGCA  gp7b-1.
725   CCAACCAGGGAGGCCCGATTTCAAAGATACTACCTGAGCA  gp7c-1.
722   CCAACCAGGGAGGCCCGATTTCAAAGATACTACCTGAGCA  gp7d-1.

      AGACAGAGGCTCAGATGGGAAACTTGGCCAGAATTTGTCT  Majority
                1050      1060      1070      1080
1041  AGACAGAGGCTCAGATGGGAAACTTGGCCAGAATTTGTCT  gphkng1815-1.
993   AGACAGAGGCTCAGATGGGAAACTTGGCCAGAATTTGTCT  gp7b-1.
765   AGACAGAGGCTCAGATGGGAAACTTGGCCAGAATTTGTCT  gp7c-1.
762   AGACAGAGGCTCAGATGGGAAACTTGGCCAGAATTTGTCT  gp7d-1.

      GATTGCGTTAATTTTCGCAAGAGATGCCAGAAATGCCAGG  Majority
                1090      1100      1110      1120
1081  GATTGCGTTAATTTTCGCAAGAGATGCCAGAAATGCCAGG  gphkng1815-1.
1033  GATTGCGTTAATTTTCGCAAGAGATGCCAGAAATGCCAGG  gp7b-1.
805   GATTGCGTTAATTTTCGCAAGAGATGCCAGAAATGCCAGG  gp7c-1.
802   GATTGCGTTAATTTTCGCAAGAGATGCCAGAAATGCCAGG  gp7d-1.
```

FIG.14G

```
      ATTATCTATCTGATGACTGCCCTAATGTGCCTGAACTATA  Majority
                1130        1140        1150        1160
1121  ATTATCTATCTGATGACTGCCCTAATGTGCCTGAACTATA  gphkng1815-1.
1073  ATTATCTATCTGATGACTGCCCTAATGTGCCTGAACTATA  gp7b-1.
845   ATTATCTATCTGATGACTGCCCTAATGTGCCTGAACTATA  gp7c-1.
842   ATTATCTATCTGATGACTGCCCTAATGTGCCTGAACTATA  gp7d-1.

      CAGAGAACTCAATGAGGCCCTCCGACTGGTCAGTAGATCC  Majority
                1170        1180        1190        1200
1161  CAGAGAACTCAATGAGGCCCTCCGACTGGTCAGTAGATCC  gphkng1815-1.
1113  CAGAGAACTCAATGAGGCCCTCCGACTGGTCAGTAGATCC  gp7b-1.
885   CAGAGAACTCAATGAGGCCCTCCGACTGGTCAGTAGATCC  gp7c-1.
882   CAGAGAACTCAATGAGGCCCTCCGACTGGTCAGTAGATCC  gp7d-1.

      AATCAGCAATACGACCAGGTGGTGCAGATGACCCAGTATC  Majority
                1210        1220        1230        1240
1201  AATCAGCAATACGACCAGGTGGTGCAGATGACCCAGTATC  gphkng1815-1.
1153  AATCAGCAATACGACCAGGTGGTGCAGATGACCCAGTATC  gp7b-1.
925   AATCAGCAATACGACCAGGTGGTGCAGATGACCCAGTATC  gp7c-1.
922   AATCAGCAATACGACCAGGTGGTGCAGATGACCCAGTATC  gp7d-1.

      ACCTGGAAGACACCACGCTTCTGATGGAGAAGATGAGAGA  Majority
                1250        1260        1270        1280
1241  ACCTGGAAGACACCACGCTTCTGATGGAGAAGATGAGAGA  gphkng1815-1.
1193  ACCTGGAAGACACCACGCTTCTGATGGAGAAGATGAGAGA  gp7b-1.
965   ACCTGGAAGACACCACGCTTCTGATGGAGAAGATGAGAGA  gp7c-1.
962   ACCTGGAAGACACCACGCTTCTGATGGAGAAGATGAGAGA  gp7d-1.
```

FIG.14H

```
      GCAGTTTGGCTGGGTTTCTGAACTGGCATACCAGTCCCCA  Majority
                1290      1300      1310      1320
1281  GCAGTTTGGCTGGGTTTCTGAACTGGCATACCAGTCCCCA  gphkng1815-1.
1233  GCAGTTTGGCTGGGTTTCTGAACTGGCATACCAGTCCCCA  gp7b-1.
1005  GCAGTTTGGCTGGGTTTCTGAACTGGCATACCAGTCCCCA  gp7c-1.
1002  GCAGTTTGGCTGGGTTTCTGAACTGGCATACCAGTCCCCA  gp7d-1.

      GGAGCTGAGGACATCTTTAATCCAGTGAAAGTAATGGTAG  Majority
                1330      1340      1350      1360
1321  GGAGCTGAGGACATCTTTAATCCAGTGAAAGTAATGGTAG  gphkng1815-1.
1273  GGAGCTGAGGACATCTTTAATCCAGTGAAAGTAATGGTAG  gp7b-1.
1045  GGAGCTGAGGACATCTTTAATCCAGTGAAAGTAATGGTAG  gp7c-1.
1042  GGAGCTGAGGACATCTTTAATCCAGTGAAAGTAATGGTAG  gp7d-1.

      CCCTAAGTGCTCATGAAGGAAATTCTTCTGATCAAGATGA  Majority
                1370      1380      1390      1400
1361  CCCTAAGTGCTCATGAAGGAAATTCTTCTGATCAAGATGA  gphkng1815-1.
1313  CCCTAAGTGCTCATGAAGGAAATTCTTCTGATCAAGATGA  gp7b-1.
1085  CCCTAAGTGCTCATGAAGGAAATTCTTCTGATCAAGATGA  gp7c-1.
1082  CCCTAAGTGCTCATGAAGGAAATTCTTCTGATCAAGATGA  gp7d-1.

      CACAGTGGTTCCTTCAAGCCTCCTGCCTTCCTCTAACTTC  Majority
                1410      1420      1430      1440
1401  CACAGTGGTTCCTTCAAGCCTCCTGCCTTCCTCTAACTTC  gphkng1815-1.
1353  CACAGTGGTTCCTTCAAGCCTCCTGCCTTCCTCTAACTTC  gp7b-1.
1125  CACAGTGGTTCCTTCAAGCCTCCTGCCTTCCTCTAACTTC  gp7c-1.
1122  CACAGTGGTTCCTTCAAGCCTCCTGCCTTCCTCTAACTTC  gp7d-1.
```

FIG.14I

```
      ACACTCAGCAGCCCTCTTGAAAAGAGTGCTGGCAACGCTA  Majority
                1450          1460          1470          1480
1441  ACACTCAGCAGCCCTCTTGAAAAGAGTGCTGGCAACGCTA  gphkng1815-1.
1393  ACACTCAGCAGCCCTCTTGAAAAGAGTGCTGGCAACGCTA  gp7b-1.
1165  ACACTCAGCAGCCCTCTTGAAAAGAGTGCTGGCAACGCTA  gp7c-1.
1162  ACACTCAGCAGCCCTCTTGAAAAGAGTGCTGGCAACGCTA  gp7d-1.

      ACTTCATTGATCACGTGGTAGAGAAGGTTCTTCAGCACTT  Majority
                1490          1500          1510          1520
1481  ACTTCATTGATCACGTGGTAGAGAAGGTTCTTCAGCACTT  gphkng1815-1.
1433  ACTTCATTGATCACGTGGTAGAGAAGGTTCTTCAGCACTT  gp7b-1.
1205  ACTTCATTGATCACGTGGTAGAGAAGGTTCTTCAGCACTT  gp7c-1.
1202  ACTTCATTGATCACGTGGTAGAGAAGGTTCTTCAGCACTT  gp7d-1.

      TAAGGAGCACTTTAAAACTTGGTAAGAAGATTTAGTCCAT  Majority
                1530          1540          1550          1560
1521  TAAGGAGCACTTTAAAACTTGGTAAGAAGATTTAGTCCAT  gphkng1815-1.
1473  TAAGGAGCACTTTAAAACTTGGTAAGAAGATTTAGTCCAT  gp7b-1.
1245  TAAGGAGCACTTTAAAACTTGGTAAGAAGATTTAGTCCAT  gp7c-1.
1242  TAAGGAGCACTTTAAAACTTGGTAAGAAGATTTAGTCCAT  gp7d-1.

      CCTATAATCAGCAAGAATTACACCTTCGGCCAAGACCTGA  Majority
                1570          1580          1590          1600
1561  CCTATAATCAGCAAGAATTACACCTTCGGCCAAGACCTGA  gphkng1815-1.
1513  CCTATAATCAGCAAGAATTACACCTTCGGCCAAGACCTGA  gp7b-1.
1285  CCTATAATCAGCAAGAATTACACCTTCGGCCAAGACCTGA  gp7c-1.
1282  CCTATAATCAGCAAGAATTACACCTTCGGCCAAGACCTGA  gp7d-1.
```

FIG.14J

```
      GAATTCTGAAAATACAAAGCAGGCTAACACAATGAACACA  Majority
                1610        1620        1630        1640
1601  GAATTCTGAAAATACAAAGCAGGCTAACACAATGAACACA  gphkng1815-1.
1553  GAATTCTGAAAATACAAAGCAGGCTAACACAATGAACACA  gp7b-1.
1325  GAATTCTGAAAATACAAAGCAGGCTAACACAATGAACACA  gp7c-1.
1322  GAATTCTGAAAATACAAAGCAGGCTAACACAATGAACACA  gp7d-1.

      GCTGCATGAAAGTTAGGTATATATTAGGAAGCACTATTGG  Majority
                1650        1660        1670        1680
1641  GCTGCATGAAAGTTAGGTATATATTAGGAAGCACTATTGG  gphkng1815-1.
1593  GCTGCATGAAAGTTAGGTATATATTAGGAAGCACTATTGG  gp7b-1.
1365  GCTGCATGAAAGTTAGGTATATATTAGGAAGCACTATTGG  gp7c-1.
1362  GCTGCATGAAAGTTAGGTATATATTAGGAAGCACTATTGG  gp7d-1.

      TTTACTTTGTTGAATGGAAGTTTAATAGCTATTCAAATTG  Majority
                1690        1700        1710        1720
1681  TTTACTTTGTTGAATGGAAGTTTAATAGCTATTCAAATTG  gphkng1815-1.
1633  TTTACTTTGTTGAATGGAAGTTTAATAGCTATTCAAATTG  gp7b-1.
1405  TTTACTTTGTTGAATGGAAGTTTAATAGCTATTCAAATTG  gp7c-1.
1402  TTTACTTTGTTGAATGGAAGTTTAATAGCTATTCAAATTG  gp7d-1.

      AGTTAATATAAAAATTTCTTCCTAAAAAGTAAAATGTACA  Majority
                1730        1740        1750        1760
1721  AGTTAATATAAAAATTTCTTCCTAAAAAGTAAAATGTACA  gphkng1815-1.
1673  AGTTAATATAAAAATTTCTTCCTAAAAAGTAAAATGTACA  gp7b-1.
1445  AGTTAATATAAAAATTTCTTCCTAAAAAGTAAAATGTACA  gp7c-1.
1442  AGTTAATATAAAAATTTCTTCCTAAAAAGTAAAATGTACA  gp7d-1.
```

FIG.14K

	TATGTAGAATATGATGCATTAGTTCTTTGTATACTAAATA	Majority
	1770 1780 1790 1800	
1761	TATGTAGAATATGATGCATTAGTTCTTTGTATACTAAATA	gphkng1815-1.
1713	TATGTAGAATATGATGCATTAGTTCTTTGTATACTAAATA	gp7b-1.
1485	TATGTAGAATATGATGCATTAGTTCTTTGTATACTAAATA	gp7c-1.
1482	TATGTAGAATATGATGCATTAGTTCTTTGTATACTAAATA	gp7d-1.

	AATACTGAGTCCCCT	Majority
	1810	
1801	AATACTGAGTCCCCT	gphkng1815-1.
1753	AATACTGAGTCCCCT	gp7b-1.
1525	AATACTGAGTCCCCT	gp7c-1.
1522	AATACTGAGTCCCCT	gp7d-1.

FIG.14L

```
                1                                                                                80
gphkng1815_aa_  MKLPLLMFPVCLLWLKDCHCAPTWKDKTAISENANSFSEAGEIDVDGEVKIALIGIKQMKIMMERREEEHSKLMKTLKKC
gp7b_aa         MKLPLLMFPVCLLWLKDCHCAPTWKDKTAISENANSFSEAGEIDVDGEVKIALIGIKQMKIMMERREEEHSKLMKTLKKC
gp7c_aa         MKLPLLMFPVCLLWLKDCHCAPTWKDKTAISENANSFSEAGEIDVDGEVKIALIGIKQMKIMMERREEEHSKLMKTLKKC
gp7d_aa         MKLPLLMFPVCLLWLKDCHCAPTWKDKTAISENANSFSEAGEIDVDGEVKIALIGIKQMKIMMERREEEHSKLMKTLKKC
                81                                                                               160
gphkng1815_aa_  KEEKQEALKLMNEVHEHLEEEESLCQVSLADSWDECRACLESNCMRFDTTCQPAWSSVKNMVEQFFRKIYQFLFPLQEND
gp7b_aa         KEEKQEALKLMNEVHEHLEEEESLCQVSLADSWDECRACLESNCMRFDTTCQPAWSSVKN.............MEND
gp7c_aa         KEEKQEALKLMNEVHEHLEEEESLCQVSLADSWDECRACLESNCMRFDTTCQPAWSSVKN.................
gp7d_aa         KEEKQEALKLMNEVHEHLEEEESLCQVSLADSWDECRACLESNCMRFDTTCQPAWSSVKN.................
                161                                                                              240
gphkng1815_aa_  RSGPVSKGVTEEDAQVSHIEHVFSQLSADVTSLFNRSLYVFKQLRREFDQAFQSYFTSGTDVTEPFFFPSLSKEPAYRAD
gp7b_aa         RSGPVSKGVTEEDAQVSHIEHVFSQLSADVTSLFNRSLYVFKQLRREFDQAFQSYFTSGTDVTEPFFFPSLSKEPAYRAD
gp7c_aa         ...............................................................MEPAYRAD
gp7d_aa         ...............................................................MPAYRAD
                241                                                                              320
gphkng1815_aa_  AEPSWAIPNVFQLLCNLSFSVYQSVSEKLITTLRATEDPPKQDKDSNQGGPISKILPEQDRGSDGKLGQNLSDCVNFRKR
gp7b_aa         AEPSWAIPNVFQLLCNLSFSVYQSVSEKLITTLRATEDPPKQDKDSNQGGPISKILPEQDRGSDGKLGQNLSDCVNFRKR
gp7c_aa         AEPSWAIPNVFQLLCNLSFSVYQSVSEKLITTLRATEDPPKQDKDSNQGGPISKILPEQDRGSDGKLGQNLSDCVNFRKR
gp7d_aa         AEPSWAIPNVFQLLCNLSFSVYQSVSEKLITTLRATEDPPKQDKDSNQGGPISKILPEQDRGSDGKLGQNLSDCVNFRKR
                321                                                                              400
gphkng1815_aa_  CQKCQDYLSDDCPNVPELYRELNEALRLVSRSNQQYDQVVQMTQYHLEDTTLLMEKMREQFGWVSELAYQSPGAEDIFNP
gp7b_aa         CQKCQDYLSDDCPNVPELYRELNEALRLVSRSNQQYDQVVQMTQYHLEDTTLLMEKMREQFGWVSELAYQSPGAEDIFNP
gp7c_aa         CQKCQDYLSDDCPNVPELYRELNEALRLVSRSNQQYDQVVQMTQYHLEDTTLLMEKMREQFGWVSELAYQSPGAEDIFNP
gp7d_aa         CQKCQDYLSDDCPNVPELYRELNEALRLVSRSNQQYDQVVQMTQYHLEDTTLLMEKMREQFGWVSELAYQSPGAEDIFNP
                401                                                                              466
gphkng1815_aa_  VKVMVALSAHEGNSSDQDDTVVPSSLLPSSNFTLSSPLEKSAGNANFIDHVVEKVLQHFKEHFKTW
gp7b_aa         VKVMVALSAHEGNSSDQDDTVVPSSLLPSSNFTLSSPLEKSAGNANFIDHVVEKVLQHFKEHFKTW
gp7c_aa         VKVMVALSAHEGNSSDQDDTVVPSSLLPSSNFTLSSPLEKSAGNANFIDHVVEKVLQHFKEHFKTW
gp7d_aa         VKVMVALSAHEGNSSDQDDTVVPSSLLPSSNFTLSSPLEKSAGNANFIDHVVEKVLQHFKEHFKTW
```

FIG.14M

```
          1                                                                  80

bhkng1    ~~~~~~~~~~~~~~~~~~~~~~~~~~GCAACCTCGTTGGTGAGAGCCTGCAGTTAGTGTCACGGCGGAAACATGAAGCCGC
bhkng2    ~~~~~~~~~~~~~~CAGAAGCTGGTGGCAACCTCGTTGGTGAGAGCCTGCAGTTAGTGTCACGGCGGAAACATGAAGCCGC
bhkng3    GTGAAGGTCCTTACAGAAGCTGGTGGCAACCTCGTTGGTGAGAGCCTGCAGTTAGTGTCACGGCGGAAACATGAAGCCGC
          81                                                                 160
bhkng1    CACTCTTGGTGTTTATTGTGTATCTGCTGCGGCTGAGAGACTGTCAGTGTGCGCCTACAGGGAAGGACCGAACTTCCATC
bhkng2    CACTCTTGGTGTTTATTGTGTATCTGCTGCGGCTGAGAGACTGTCAGTGTGCGCCTACAGGGAAGGACCGAACTTCCATC
bhkng3    CAATCTTGGTGTTTATCGTGTATCTGCTGCAGCTGAGAGACTGTCAGTGTGCGCCTACAGGGAAGGACCGAACTTCCATC
          161                                                                240
bhkng1    CGTGAAGACCCGAAGGGTTTTTCCAAGGCTGGGGAGATAGACGTAGATGAAGAGGTGAAGAAGGCTTTGATTGGCATGAA
bhkng2    CGTGAAGACCCGAAGGGTTTTTCCAAGGCTGGGGAGATAGACGTAGATGAAGAGGTGAAGAAGGCTTTGATTGGCATGAA
bhkng3    CGTGAAGACCCGAAGGGTTTTTCCAAGGCTGGGGAGATAGACGTAGATGAAGAGGTGAAGAAGGCTTTGATTGGCATGAA
          241                                                                320
bhkng1    GCAGATGAAAATCCTGATGGAAAGAAGAGAGGAGGAACATAGCAAACTAATGAGAACACTGAAGAAATGCAGAGAAGAAA
bhkng2    GCAGATGAAAATCCTGATGGAAAGAAGAGAGGAGGAACATAGCAAACTAATGAGAACACTGAAGAAATGCAGAGAAGAAA
bhkng3    GCAGATGAAAATCCTGATGGAAAGAAGAGAGGAGGAACATAGCAAACTAATGAGAACACTGAAGAAATGCAGAGAAGAAA
          321                                                                400
bhkng1    AGCAGGAGGCCCTGAAGCTTATGAATGAAGTTCAAGAACATCTAGAAGAGGAAGAAAGGCTATGCCAGGTGTCTCTGATG
bhkng2    AGCAGGAGGCCCTGAAGCTTATGAATGAAGTTCAAGAACATCTAGAAGAGGAAGAAAGGCTATGCCAGGTGTCTCTGATG
bhkng3    AGCAGGAGGCCCTGAAGCTTATGAATGAAGTTCAAGAACATCTAGAAGAGGAAGAAAGGCTATGCCAGGTGTCTCTGATG
          401                                                                480
bhkng1    GGTTCCTGGGACGAATGCAAATCTTGCCTGGAAAGTGACTGCATGAGATTTTATACAACCTGCCAAAGCAGTTGGTCCTC
bhkng2    GGTTCCTGGGACGAATGCAAATCTTGCCTGGAAAGTGACTGCATGAGATTTTATACAACCTGCCAAAGCAGTTGGTCCTC
bhkng3    GGTTCCTGGGACGAATGCAAATCTTGCCTGGAAAGTGACTGCATGAGATTTTATACAACCTGCCAAAGCAGTTGGTCCTC
          481                                                                560
bhkng1    TATGAAATCCACGATTGAACGGGTTTTCCGGAAGATATATCAGTTTCTCTTTCCTTTCCATGAAGACGATGAAAAAGAGC
bhkng2    TATGAAATCCACGATTGAACGGGTTTTCCGGAAGATATATCAGTTTCTCTTTCCTTTCCATGAAGACGATGAAAAAGAGC
bhkng3    TATGAAATCCACGATTGAACGGGTTTTCCGGAAGATATATCAGTTTCTCTTTCCTTTCCATGAAGACGATGAAAAAGAGC
```

FIG.15A

```
        561                                                                640
bhkng1  TTCCTGTTGGTGAGAAGTTCACTGAGGAAGATGTACAGCTGATGCAGATAGAGAATGTGTTCAGCCAGCTGACCGTGGAT
bhkng2  TTCCTGTTGGTGAGAAGTTCACTGAGGAAGATGTACAGCTGATGCAGATAGAGAATGTGTTCAGCCAGCTGACCGTGGAT
bhkng3  TTCCTGTTGGTGAGAAGTTCACTGAGGAAGATGTACAGCTGATGCAGATAGAGAATGTGTTCAGCCAGCTGACCGTGGAC
        641                                                                720
bhkng1  GTGGGATTTCTCTATAACATGAGCTTTCACGTCTTCAAACAGATGCAGCAAGAATTTGACCTGGCTTTTCAATCATACTT
bhkng2  GTGGGATTTCTCTATAACATGAGCTTTCACGTCTTCAAACAGATGCAGCAAGAATTTGACCTGGCTTTTCAATCATACTT
bhkng3  GTGGGATTTCTCTATAACATGAGCTTTCACGTCTTCAAACAGATGCAGCAAGAATTTGACCTGGCTTTTCAATCATACTT
        721                                                                800
bhkng1  TATGTCAGACACAGACTCCATGGAGCCTTACTTTTTTTCCAGCTTTTTTCCAAAGAGCCAGCAAAAAAAGCACATCCTATGC
bhkng2  TATGTCAGACACAGACTCCATGGAGCCTTACTTTTTTTCCAGCTTTTTTCCAAAGAGCCAGCAAAAAAAGCACATCCTATGC
bhkng3  TATGTCAGACACAGACTCCATGGAGCCTTACTTTTTTTCCAGCTTTTTTCCAAAGAGCCAGCAAAAAAAGCACATCCTATGC
        801                                                                880
bhkng1  AGAGTTGGGACATTCCCAGCTTCTTCCAGCTGTTTTGTAATTTCAGCCTCTCTGTTTATCAAAGTGTCAGCGCAACAGTT
bhkng2  AGAGTTGGGACATTCCCAGCTTCTTCCAGCTGTTTTGTAATTTCAGCCTCTCTGTTTATCAAAGTGTCAGCGCAACAGTT
bhkng3  AGAGTTGGGACATTCCCAGCTTCTTCCAGCTGTTTTGTAATTTCAGCCTCTCTGTTTATCAAAGTGTCAGCGCAACAGTT
        881                                                                960
bhkng1  ACAGAGATGCTGAAGGCCATTGAGGACTTATCCAAACAAGACAAAGATTCTGCCCACGGTGGACCGAGTTCCACGACGTG
bhkng2  ACAGAGATGCTGAAGGCCATTGAGGACTTATCCAAACAAGACAAAGATTCTGCCCACGGTGGACCGAGTTCCACGACGTG
bhkng3  ACAGAGATGCTGAAGGCCATTGAGGACTTATCCAAACAAGACAAAGATTCTGCCCACGGTGGACCGAGTTCCACGACGTG
        961                                                               1040
bhkng1  GCCTGTGCGGGGCAGAGGGCTGTGTGGAGAACCTGGCCAGAACTCGTCCGAATGTCTCCAATTTCATGCAAGATGCCAGA
bhkng2  GCCTGTGCGGGGCAGAGGGCTGTGTGGAGAACCTGGCCAGAACTCGTCCGAATGTCTCCAATTTCATGCAAGATGCCAGA
bhkng3  GCCTGTGCGGGGCAGAGGGCTGTGTGGAGAACCTGGCCAGAACTCGTCCGAATGTCTCCAATTTCATGCAAGATGCCAGA
        1041                                                              1120
bhkng1  AATGTCAGGATTACCTATGGGCAGACTGCCCTGCTGTTCCTGAACTATACACAAAGGCGGATGAGGCCCTTGAGTTGGTC
bhkng2  AATGTCAGGATTACCTATGGGCAGACTGCCCTGCTGTTCCTGAACTATACACAAAGGCGGATGAGGCCCTTGAGTTGGTC
bhkng3  AATGTCAGGATTACCTATGGGCAGACTGCCCTGCTGTTCCTGAACTATACACAAAGGCGGATGAGGCCCTTGAGTTGGTC
```

FIG.15B

```
        1121                                                                1200
bhkng1  AACATATCCAATCAGCAGTATGCCCAGGTACTCCAGATGACCCAGCATCACTTGGAGGACACCACGTATCTGATGGAGAA
bhkng2  AACATATCCAATCAGCAGTATGCCCAGGTACTCCAGATGACCCAGCATCACTTGGAGGACACCACGTATCTGATGGAGAA
bhkng3  AACATATCCAATCAGCAGTATGCCCAGGTACTCCAGATGACCCAGCATCACTTGGAGGACACCACGTATCTGATGGAGAA
        1201                                                                1280
bhkng1  GATGAGAGAGCAGTTTGGTTGGGTAACAGAGCTGGCCAGCCAGACCCCAGGAAGCGAGAACATCTTCAGTTTCATAAAGG
bhkng2  GATGAGAGAGCAGTTTGGTTGGGTAACAGAGCTGGCCAGCCAGACCCCAGGAAGCGAGAACATCTTCAGTTTCATAAAGG
bhkng3  GATGAGAGAGCAGTTTGGTTGGGTAACAGAGCTGGCCAGCCAGACCCCAGGAAGCGAGAACATCTTCAGTTTCATAAAGG
        1281                                                                1360
bhkng1  TAGTTCCAGGTGTTCACGAAGGAAATTTCTCCAAACAAGATGAAAAGATGATAGACATAAGCATTCTGCCTTCCTCTAAT
bhkng2  TAGTTCCAGGTGTTCACGAAGGAAATTTCTCCAAACAAGATGAAAAGATGATAGACATAAGCATTCTGCCTTCCTCTAAT
bhkng3  TAGTTCCAGGTGTTCACGAAGGAAATTTCTCCAAACAAGATGAAAAGATGATAGACATAAGCATTCTGCCTTCCTCTAAT
        1361                                                                1440
bhkng1  TTCACACTCACCATCCCTCTTGAAGAAAGTGCTGAGAGTTCCGACTTCATTAGCTACATGCTGGCCAAAGCTGTACAGCA
bhkng2  TTCACACTCACCATCCCTCTTGAAGAAAGTGCTGAGAGTTCCGACTTCATTAGCTACATGCTGGCCAAAGCTGTACAGCA
bhkng3  TTCACACTCACCATCCCTCTTGAAGAAAGTGCTGAGAGTTCCGACTTCATTAGCTACATGCTGGCCAAAGCTGTACAGCA
        1441                                                                1530
bhkng1  TTTTAAGGAACATTTTAAATCTTGGTAAGCAGAGTATTTGATTAGGGACGTTTGCTGATAGGAATAGATGGTTCTTAAAA
bhkng2  TTTTAAGGAACATTTTAAATCTTGGTAAGCAGAGTATTTGATTAGGGACGTTTGCTGATAGGAATAGATGGTTCTTAAAA
bhkng3  TTTTAAGGAACATTTTAAATCTTGGTAAGCAGAGTATTTGATTAGGGACGTTTGCTGATAGGAATAGATGGTTCTTAAAA
        1521                                                                1600
bhkng1  GGGAAAAATGACAAAACTAGCTTTTGAATACCTTGAAAACGTATTCAACCTCATTAATAATCAAAGGCATGAAAACTAAG
bhkng2  GGGAAAAATGACAAAACTAGCTTTTGAATACCTTGAAAACGTATTCAACCTCATTAATAATCAAAGGCATGAAAACTAAG
bhkng3  GGGAAAAATGACAAAACTAGCTTTTGAATACCTTGAAAACGTATTCAACCTCATTAATAATCAAAGGCATGAAAACTAAG
        1601                                                                1680
bhkng1  ACAAGTTAGCAGTTTTTACCTATTGAATTTTCAAATTAAAAAAAAAAATCCTGATAGAATGCAATGAAATGAGAATTCTT
bhkng2  ACAAGTTAGCAGTTTTTACCTATTGAATTTTCAAATTAAAAAAAAAAATCCTGATAGAATGCAATGAAATGAGAATTCTT
bhkng3  ACAAGTTAGCAGTTTTTACCTATTGAATTTTCAAATTAAAAAAAAAA.TCCTGATAGAATGCAATGAAATGAGAATTCTT
```

FIG.15C

```
        1681                                                                1760
bhkng1  ATATGTGATTGCCAGAAACAAACTGGTTTTGTCTTTTTGAAAAGTTATTCAATTATACATATCAAGAGTCATCAAATTTC
bhkng2  ATATGTGATTGCCAGAAACAAACTGGTTTTGTCTTTTTGAAAAGTTATTCAATTATACATATCAAGAGTCATCAAATTTC
bhkng3  ATATGTGATTGCCAGAAACAAACTGGTTTTGTCTTTTTGAAAAGTTATTCAATTATACATATCAAGAGTCATCAAATTTC
        1761                                                                1840
bhkng1  TTTTTAATATAATAATTCCACTTCTGGAATCAATCCAAAGGAGTAAATCTAAAATTGAATTGAAGTTCCCACCCCAAGAT
bhkng2  TTTTTAATATAATAATTCCACTTCTGGAATCAATCCAAAGGAGTAAATCTAAAATTGAATTGAAGTTCCCACCCCAAGAT
bhkng3  TTTTTAATATAATAATTCCACTTCTGGAATCAATCCAAAGGAGTAAATCTAAAATTGAATTGAAGTTCCCACCCCAAGAT
        1841                                                                1920
bhkng1  CAATATTTGCAAATTATTTAAAATAGTAAACTGTTAAAAACTGAATGTCATCTGAATGTCTAAAAACCAGAAATGGTTAA
bhkng2  CAATATTTGCAAATTATTTAAAATAGTAAACTGTTAAAAACTGAATGTCATCTGAATGTCTAAAAACCAGAAATGGTTAA
bhkng3  CAATATTTGCAAATTATTTAAAATAGTAAACTGTTAAAAACTGAATGTCATCTGAATGTCTAAAAACCAGAAATGGTTAA
        1921                                                                2000
bhkng1  AAGCTGTGGCTAAATATGCTCCAAATATCTTATAAAACCATTAAAAATATTTATAAAATTTAAATCATGACATGACATCT
bhkng2  AAGCTGTGGCTAAATATGCTCCAAATATCTTATAAAACCATTAAAAATATTTATAAAATTTAAATCATGACATGACATCT
bhkng3  AAGCTGTGGCTAAATATGCTCCAAATATCTTATAAAACCATTAAAAATATTTATAAAATTTAAATCATGACATGACATCT
        2001                                                                2080
bhkng1  GCTGGAACAAGAGTTTATTCTAAGCCTATCTATAAGGCAAATATTATTATTACTATCTTCCAGAAAAGAAACTTGAGACT
bhkng2  GCTGGAACAAGAGTTTATTCTAAGCCTATCTATAAGGCAAATATTATTATTACTATCTTCCAGAAAAGAAACTTGAGACT
bhkng3  GCTGGAACAAGAGTTTATTCTAAGCCTATCTATAAGGCAAATATTATTATTACTATCTTCCAGAAAAGAAACTTGAGACT
        2081                                                                2160
bhkng1  CAGGGTCCAAGTGTTAGTTGCTCAGTCATGTCTGACTCTTTGGGACCCCTTGGACTGTAGCCCACCAGGCTCCTCTGTCC
bhkng2  CAGGGTCCAAGTGTTAGTTGCTCAGTCATGTCTGACTCTTTGGGACCCCTTGGACTGTAGCCCACCAGGCTCCTCTGTCC
bhkng3  CAGGGTCCAAGTGTTAGTTGCTCAGTCATGTCTGACTCTTTGGGACCCCTTGGACTGTAGCCCACCAGGCTCCTCTGTCC
        2161                                                                2240
bhkng1  GTGGGATTCTTCAGACAGGAATACTGGGGCAGGTTGCTATTTCCTTCTCCAGGAAATCTTCCCTATCCAGGGATGGAACC
bhkng2  ATGGGATTCTTCAGACAGGAATACTGGGGCAGGTTGCTATTTCCTTCTCCAGGAAATCTTCCCTATCCAGGGATGGAACC
bhkng3  ATGGGATTCTTCAGACAGGAATACTGGGGCAGGTTGCTATTTCCTTCTCCAGGAAATCTTCCCTATCCAGGGATGGAACC
```

FIG.15D

```
        2241                                                                2320
bhkng1  CAGGTCTCCTGCATTGCAGGTAGATGCTTTACTATCTGAGCAACCAAATGAATTACTCAAGTCAGTAGGGGGTAGAGGCA
bhkng2  CAGGTCTCCTGCATTGCAGGTAGATGCTTTACTATCTGAGCAACCAAATGAATTACTCAAGTCAGTAGGGGGTAGAGGCA
bhkng3  CAGGTCTCCTGCATTGCAGGTAGATGCTTTACTATCTGAGCAACCAAATGAATTACTCAAGTCAGTAGGGGGTAGAGGCA
        2321                                                                2400
bhkng1  AATTTTAACTTAGTTTTCTCTGAATCATAATTGCCACATTAAACTGGTTCCTGTTGGGACATTTGGTTGAAAAAAATAAA
bhkng2  AATTTTAACTTAGTTTTCTCTGAATCATAATTGCCACATTAAACTGGTTCCTGTTGGGACATTTGGTTGAAAAAAATAAA
bhkng3  AATTTTAACTTAGTTTTCTCTGAATCATAATTGCCACATTAAACTGGTTCCTGTTGGGACATTTGGTTGAAAAAAATAAA
        2401                                                                2480
bhkng1  GTGAAAAATGAGTATAAAACTCTATAAATGTAATGATCAAAACGAAAAAAAAATCTACAATCTGCATTAAAAATAAAAAGG
bhkng2  GTGAAAAATGAGTATAAAACTCTATAAATGTAATGATCAAAACGAAAAAAAAATCTACAATCTGCATTAAAAATAAAAAGG
bhkng3  GTGAAAAATGAGTATAAAACTCTATAAATGTAATGATCAAAACGAAAAAAAAATCTACAATCTGCATTAAAAATAAAAAGG
        2481                                                                2560
bhkng1  GTTGGCAGG.....................................................................
bhkng2  GTTGGCAGGAATTACGGTTGGAAATGGATGATTTTTTTTAACCTTTTCATCTTTTGATATTTTACAATTTTCTATAATGA
bhkng3  GTTGGCAGG.....................................................................
        2561                                                                2640
bhkng1  .....................................................................
bhkng2  ATAAATAATTTTGAGATTTCAAATTAGAAGATATGTTGCTAAAATAGCTAGGTAAATGTAGATTGAACACTGTATCAATG
bhkng3  .....................................................................
        2641                                                                2720
bhkng1  .....................................................................
bhkng2  TGTTCTCATCTTTAAACTTTAGTATAAGTACTTCTATTCCATGGTAATCCTACAGTAAGACGAAATGTAAATCTGTTCGG
bhkng3  .....................................................................
        2721                                                                2800
bhkng1  .....................................................................
bhkng2  TCTACAGGAAAAAACAACTAAATGACATTTCAGACGTACATTACCATCTCTGTTAGGATAATCTTCTGAATTAATGGCAC
bhkng3  .....................................................................
```

FIG.15E

```
          2801                                                                2880
bhkng1    ....................................................................
bhkng2    ATTAGAACTGTACATAGTATTCTCCTTTGGTAAAATGGTCAATCTTAAAGAAGCATTAAATGTTAATTCTAAGTTATTAC
bhkng3    ....................................................................
          2881                                                                2960
bhkng1    ....................................................................
bhkng2    TCATAAGGGACCTTGTAGGTAGGTCCCTATCAATGTATAATTAAGCTGGGTATTTCTAGATTCGCTGCCTCTCCCTTTAT
bhkng3    ....................................................................
          2961                                                                3029
bhkng1    ....................................................................
bhkng2    CTCTGAATGTTGGAGAGGTTGTTGGTCATCAATCAACCAATATCTTTTTAGCATCTTCTAAGTGAAGGC
bhkng3    ....................................................................
```

FIG.15F

```
                1                                                                80
hmhkng_aa       MKIKAEKNEGPSRSWWQLHWGDIANNSGNMKPPLLVFIVCLLWLKDSHCAPTWKDKTAISENLKSFSEVGEIDADEEVKK
bhkng1_aa       ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~MKPPLLVFIVYLLRLRDCQCAPTGKDRTSIREDPKGFSKAGEIDVDEEVKK
phkng1815_aa_   ~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~MKLPLLMFPVCLLWLKDCHCAPTWKDKTAISENANSFSEAGEIDVDGEVKI
               81                                                               160
hmhkng_aa       ALTGIKQMKIMMERKEKEHTNLMSTLKKCREEKQEALKLLNEVQEHLEEEERLCRESLADSWGECRSCLENNCMRIYTTC
bhkng1_aa       ALIGMKQMKILMERREEEHSKLMRTLKKCREEKQEALKLMNEVQEHLEEEERLCQVSLMGSWDECKSCLESDCMRFYTTC
gphkng1815_aa_  ALIGIKQMKIMMERREEEHSKLMKTLKKCKEEKQEALKLMNEVHEHLEEEESLCQVSLADSWDECRACLESNCMRFDTTC
              161                                                               240
hmhkng_aa       QPSWSSVKNKIERFFRKIYQFLFPFHEDNEKDLPISEKLIEEDAQLTQMEDVFSQLTVDVNSLFNRSFNVFRQMQQEFDQ
bhkng1_aa       QSSWSSMKSTIERVFRKIYQFLFPFHEDDEKELPVGEKFTEEDVQLMQIENVFSQLTVDVGFLYNMSFHVFKQMQQEFDL
gphkng1815_aa_  QPAWSSVKNMVEQFFRKIYQFLFPLQE.NDRSGPVSKGVTEEDAQVSHIEHVFSQLSADVTSLFNRSLYVFKQLRREFDQ
              241                                                               320
hmhkng_aa       TFQSHFISDTDLTEPYFFPAFSKEPMTKADLEQCWDIPNFFQLFCNFSVSIYESVSETITKMLKAIEDLPKQDKAPDHGG
bhkng1_aa       AFQSYFMSDTDSMEPYFFPAFSKEPAKKAHPMQSWDIPSFFQLFCNFSLSVYQSVSATVTEMLKAIEDLSKQDKDSAHGG
gphkng1815_aa_  AFQSYFTSGTDVTEPFFFPSLSKEPAYRADAEPSWAIPNVFQLLCNLSFSVYQSVSEKLITTLRATEDPPKQDKDSNQGG
              321                                                               400
hmhkng_aa       LISKMLPGQDRGLCGELDQNLSRCFKFHEKCQKCQAHLSEDCPDVPALHTELDEAIRLVNVSNQQYGQILQMTRKHLEDT
bhkng1_aa       PSSTTWPVRGRGLCGEPGQNSSECLQFHARCQKCQDYLWADCPAVPELYTKADEALELVNISNQQYAQVLQMTQHHLEDT
gphkng1815_aa_  PISKILPEQDRGSDGKLGQNLSDCVNFRKRCQKCQDYLSDDCPNVPELYRELNEALRLVSRSNQQYDQVVQMTQYHLEDT
              401                                                               480
hmhkng_aa       AYLVEKMRGQFGWVSELANQAPETEIIFNSIQVVPRI..HEGNISKQDETMMTDLSILPSSNFTLKIPLEESAESSNFIG
bhkng1_aa       TYLMEKMREQFGWVTELASQTPGSENIFSFIKVVPGV..HEGNFSKQDE.KMIDISILPSSNFTLTIPLEESAESSDFIS
gphkng1815_aa_  TLLMEKMREQFGWVSELAYQSPGAEDIFNPVKVMVALSAHEGNSSDQDD.TVVPSSLLPSSNFTLSSPLEKSAGNANFID
              481                                                               497
hmhkng_aa       YVVAKALQHFKEHFKTW
bhkng1_aa       YMLAKAVQHFKEHFKSW
gphkng1815_aa_  HVVEKVLQHFKEHFKTW
```

FIG.16

```
mature HKNG       ----------------------------------------------APTWKDKTAIS
HKNG1-V1-IPF3     ------------------MRTWDYSNSGNMKPPLLVFIVCLLWLKDSHCAPTWKDKTAIS
HKNG1/1-V1-IPF2   ----------------------------MKPPLLVFIVCLLWLKDSHCAPTWKDKTAIS
HKNG1-IPF1        MKIKAEKNEGPSRSWWQLHWGDIANNSGNMKPPLLVFIVCLLWLKDSHCAPTWKDKTAIS
                                                           ***********

mature HKNG       ENLKSFSEVGEIDADEEVKKALTGIKQMKIMMERKEKEHTNLMSTLKKCREEKQEALKLL
HKNG1-V1-IPF3     ENLKSFSEVGEIDADEEVKKALTGIKQMKIMMERKEKEHTNLMSTLKKCREEKQEALKLL
HKNG1/1-V1-IPF2   ENLKSFSEVGEIDADEEVKKALTGIKQMKIMMERKEKEHTNLMSTLKKCREEKQEALKLL
HKNG1-IPF1        ENLKSFSEVGEIDADEEVKKALTGIKQMKIMMERKEKEHTNLMSTLKKCREEKQEALKLL
                  ************************************************************

mature HKNG       NEVQEHLEEEERLCRESLADSWGECRSCLENNCMRIYTTCQPSWSSVKNKIERFFRKIYQ
HKNG1-V1-IPF3     NEVQEHLEEEERLCRESLADSWGECRSCLENNCMRIYTTCQPSWSSVKNKIERFFRKIYQ
HKNG1/1-V1-IPF2   NEVQEHLEEEERLCRESLADSWGECRSCLENNCMRIYTTCQPSWSSVKNKIERFFRKIYQ
HKNG1-IPF1        NEVQEHLEEEERLCRESLADSWGECRSCLENNCMRIYTTCQPSWSSVKNKIERFFRKIYQ
                  ************************************************************

mature HKNG       FLFPFHEDNEKDLPISEKLIEEDAQLTQMEDVFSQLTVDVNSLFNRSFNVFRQMQQEFDQ
HKNG1-V1-IPF3     FLFPFHEDNEKDLPISEKLIEEDAQLTQMEDVFSQLTVDVNSLFNRSFNVFRQMQQEFDQ
HKNG1/1-V1-IPF2   FLFPFHEDNEKDLPISEKLIEEDAQLTQMEDVFSQLTVDVNSLFNRSFNVFRQMQQEFDQ
HKNG1-IPF1        FLFPFHEDNEKDLPISEKLIEEDAQLTQMEDVFSQLTVDVNSLFNRSFNVFRQMQQEFDQ
                  ************************************************************

mature HKNG       TFQSHFISDTDLTEPYFFPAFSKEPMTKADLEQCWDIPNFFQLFCNFSVSIYESVSETIT
HKNG1-V1-IPF3     TFQSHFISDTDLTEPYFFPAFSKEPMTKADLEQCWDIPNFFQLFCNFSVSIYESVSETIT
HKNG1/1-V1-IPF2   TFQSHFISDTDLTEPYFFPAFSKEPMTKADLEQCWDIPNFFQLFCNFSVSIYESVSETIT
HKNG1-IPF1        TFQSHFISDTDLTEPYFFPAFSKEPMTKADLEQCWDIPNFFQLFCNFSVSIYESVSETIT
                  ************************************************************
```

FIG.17A

```
mature HKNG       KMLKAIEDLPKQDKAPDHGGLISKMLPGQDRGLCGELDQNLSRCFKFHEKCQKCQAHLSE
HKNG1-V1-IPF3     KMLKAIEDLPKQDKAPDHGGLISKMLPGQDRGLCGELDQNLSRCFKFHEKCQKCQAHLSE
HKNG1/1-V1-IPF2   KMLKAIEDLPKQDKAPDHGGLISKMLPGQDRGLCGELDQNLSRCFKFHEKCQKCQAHLSE
HKNG1-IPF1        KMLKAIEDLPKQDKAPDHGGLISKMLPGQDRGLCGELDQNLSRCFKFHEKCQKCQAHLSE
                  ************************************************************

mature HKNG       DCPDVPALHTELDEAIRLVNVSNQQYGQILQMTRKHLEDTAYLVEKMRGQFGWVSELANQ
HKNG1-V1-IPF3     DCPDVPALHTELDEAIRLVNVSNQQYGQILQMTRKHLEDTAYLVEKMRGQFGWVSELANQ
HKNG1/1-V1-IPF2   DCPDVPALHTELDEAIRLVNVSNQQYGQILQMTRKHLEDTAYLVEKMRGQFGWVSELANQ
HKNG1-IPF1        DCPDVPALHTELDEAIRLVNVSNQQYGQILQMTRKHLEDTAYLVEKMRGQFGWVSELANQ
                  ************************************************************

mature HKNG       APETEIIFNSIQVVPRIHEGNISKQDETMMTDLSILPSSNFTLKIPLEESAESSNFIGYV
HKNG1-V1-IPF3     APETEIIFNSIQVVPRIHEGNISKQDETMMTDLSILPSSNFTLKIPLEESAESSNFIGYV
HKNG1/1-V1-IPF2   APETEIIFNSIQVVPRIHEGNISKQDETMMTDLSILPSSNFTLKIPLEESAESSNFIGYV
HKNG1-IPF1        APETEIIFNSIQVVPRIHEGNISKQDETMMTDLSILPSSNFTLKIPLEESAESSNFIGYV
                  ************************************************************

mature HKNG       VAKALQHFKEHFKTW
HKNG1-V1-IPF3     VAKALQHFKEHFKTW
HKNG1/1-V1-IPF2   VAKALQHFKEHFKTW
HKNG1-IPF1        VAKALQHFKEHFKTW
                  ***************
```

FIG.17B

```
    R   H   L   Q   A   R   A   A   G   L   V   S   T   L   E   V   A   D   T    19
TG CGT CAC CTG CAG GCC CGG GCC GCG GGG TTG GTT TCC ACC CTG GAG GTT GCT GAC ACC    57

 L   C   P   R   L   T   S   S   R   W   H   R   R   L   Q   G   A   A   L   K    39
CTG TGC CCT CGG CTG ACT TCC AGC CGG TGG CAC AGA CGC CTC CAG GGG GCA GCA CTC AAG  117

 R   I   L   G   M   T   E   L   R   P   S   L   L   P   G   W   S   S   V   A    59
CGC ATC TTA GGA ATG ACA GAG TTG CGT CCC TCT CTG TTG CCA GGC TGG AGT TCA GTG GCA  177

 C   S   *   L   T   E   A   S   N   S   W   V   Q   V   T   L   P   P   Q   P    79
TGT TCT TAG CTC ACT GAA GCC TCA AAT TCC TGG GTT CAA GTG ACC CTC CCA CCT CAG CCC  237

 H   E   D   L   G   L   Q   D   T   A   K   S   L   T   R   M   K   I   K   A    99
CAT GAG GAC CTG GGA CTA CAG GAC ACA GCT AAA TCC CTG ACA CGG ATG AAA ATT AAA GCA  297

 E   K   N   E   G   P   S   R   S   W   W   Q   L   H   W   G   D   I   A   N   119
GAG AAA AAC GAA GGT CCT TCC AGA AGC TGG TGG CAA CTT CAC TGG GGA GAT ATT GCA AAT  357

 N   S   G   N   M   K   P   P   L   L   V   F   I   V   C   L   L   W   L   K   139
AAC AGC GGG AAC ATG AAG CCG CCA CTC TTG GTG TTT ATT GTG TGT CTG CTG TGG TTG AAA  417

 D   S   H   C   A   P   T   W   K   D   K   T   A   I   S   E   N   L   K   S   159
GAC AGT CAC TGC GCA CCC ACT TGG AAG GAC AAA ACT GCT ATC AGT GAA AAC CTG AAG AGT  477

 F   S   E   V   G   E   I   D   A   D   E   E   V   K   K   A   L   T   G   I   179
TTT TCT GAG GTG GGG GAG ATA GAT GCA GAT GAA GAG GTG AAG AAG GCT TTG ACT GGT ATT  537

 K   Q   M   K   I   M   M   E   R   K   E   K   E   H   T   N   L   M   S   T   199
AAG CAA ATG AAA ATC ATG ATG GAA AGA AAA GAG AAG GAA CAC ACC AAT CTA ATG AGC ACC  597
```

FIG.18A

```
 L  K  K  C  R  E  E  K  Q  E  A  L  K  L  L  N  E  V  Q  E  219
CTG AAG AAA TGC AGA GAA GAA AAG CAG GAG GCC CTG AAA CTT CTG AAT GAA GTT CAA GAA  657

 H  L  E  E  E  E  R  L  C  R  E  S  L  A  D  S  W  G  E  C  239
CAT CTG GAG GAA GAA GAA AGG CTA TGC CGG GAG TCT TTG GCA GAT TCC TGG GGT GAA TGC  717

 R  S  C  L  E  N  N  C  M  R  I  Y  T  T  C  Q  P  S  W  S  259
AGG TCT TGC CTG GAA AAT AAC TGC ATG AGA ATT TAT ACA ACC TGC CAA CCT AGC TGG TCC  777

 S  V  K  N  K  L  L  T  T  E  A  *  F  Q  R  C  Y  L  G  R  279
TCT GTG AAA AAT AAG CTC CTG ACC ACG GAG GCC TGA TTT CAA AGA TGT TAC XTG GGC AGG  837

 T  E  D  C  V  G  N  L  T  R  I  C  Q  D  V  S  N  F  M  K  299
ACA GAG GAC TGT GTG GGG AAC TTG ACC AGA ATT TGT CAA GAT GTT TCA AAT TTC ATG AAA  897
 N  A  K  N  V  R  L  T  Y  L  K  T  V  L  M  Y  L  L  C  T  319
AAT GCC AAA AAT GTC AGG CTC ACC TAT CTG AAG ACT GTC CTG ATG TAC CTG CTC TGC ACA  957

 Q  N  *  T  R  R  S  G  W  S  M  Y  P  I  S  S  M  A  R  F  339
CAG AAT TAG ACG AGG CGA TCA GGT TGG TCA ATG TAT CCA ATC AGC AGT ATG GCC AGA TTC 1017

 S  R  *  P  G  S  T  W  R  T  P  P  I  W  W  R  R  *  E  G  359
TCC AGA TGA CCC GGA AGC ACT TGG AGG ACA CCG CCT ATC TGG TGG AGA AGA TGA GAG GGC 1077

 N  L  A  G  C  L  N  W  Q  T  R  P  Q  K  Q  R  S  S  L  I  379
AAT TTG GCT GGG TGT CTG AAC TGG CAA ACC AGG CCC CAG AAA CAG AGA TCA TCT TTA ATT 1137

 Q  Y  R  *  F  Q  G  F  M  K  E  I  F  P  N  K  M  K  Q  *  399
CAA TAC AGG TAG TTC CAA GGA TTC ATG AAG GAA ATA TTT CCA AAC AAG ATG AAA CAA TGA 1197
```

FIG.18B

```
 *   Q   T   *   A   F   C   L   P   L   I   S   H   S   R   S   L   L   K   K  419
TGA CAG ACT TAA GCA TTC TGC CTT CCT CTA ATT TCA CAC TCA AGA TCC CTC TTG AAG AAA 1257

 V   L   R   V   L   T   S   L   A   T   *   W   Q   K   L   Y   S   I   L   R  439
GTG CTG AGA GTT CTA ACT TCA TTG GCT ACG TAG TGG CAA AAG CTC TAC AGC ATT TTA AGG 1317

 N   I   L   K   P   G   K   K   I   *   C   I   L   Y   P   V   S   R   I   I  459
AAC ATT TTA AAA CCT GGT AAG AAG ATC TAA TGC ATC CTA TAT CCA GTA AGT AGA ATT ATC 1377

 S   S   S   G   T   W   K   S   *   N   K   K   G   *   C   N   K   H   S   C  479
TCT TCA TCT GGG ACC TGG AAA TCC TGA AAT AAA AAA GGA TAA TGC AAT AAA CAC AGT TGC 1437

 R   K   V   C   *   L   Y   T   M   K   Y   S   *   F   T   Y   V   E   W   L  499
AGG AAA GTA TGT TAG CTA TAT ACT ATG AAG TAC TCT TAG TTT ACT TAT GTT GAA TGG CTT 1497

 S   Y   *   Y   S   N   *   V   K   M   K   I   P   P   *   K   I   K   R   N  519
AGC TAT TAA TAC TCA AAT TGA GTT AAA ATG AAA ATT CCT CCT TAA AAA ATC AAA CGT AAT 1557

 M   Y   Y   I   S   W   Y   I   S   S   S   L   Y   I   E   *   I   L   N   H  539
ATG TAT TAC ATT TCA TGG TAC ATT AGT AGT TCT TTG TAT ATT GAA TAA ATA CTA AAT CAC 1617

 L                                                                              540
CTA                                                                             1620
```

FIG.18C

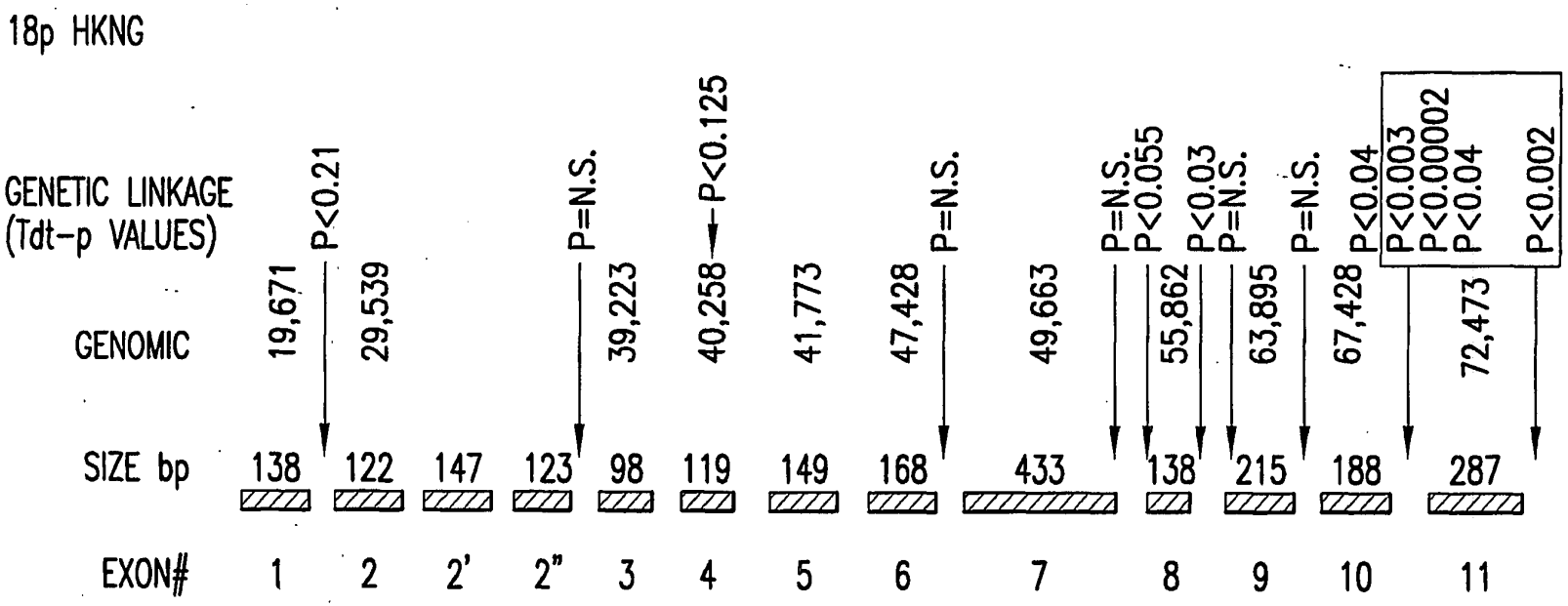
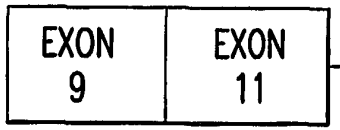

18p HKNG
GENETIC LINKAGE
(Tdt-p VALUES)
GENOMIC
SIZE bp
EXON#
19,671
P<0.21
29,539
P=N.S.
39,223
40,258
P<0.125
41,773
47,428
P=N.S.
49,663
P=N.S.
P<0.055
55,862
P<0.03
P=N.S.
63,895
P=N.S.
67,428
P<0.04
P<0.003
P<0.00002
P<0.04
72,473
P<0.002
138
122
147
123
98
119
149
168
433
138
215
188
287
1
2
2'
2"
3
4
5
6
7
8
9
10
11

FIG.19

FIG. 20A

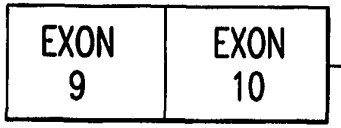
WILD TYPE

FIG. 20B



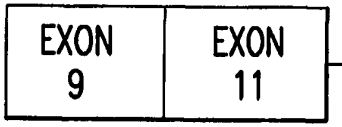
DELTA 10–DELETED FOR BASES 1581–1768

FIG. 20C



EXTENDING 125bp INTO INTRON 10

FIG. 20D



DELTA 10+; LONGEST EXTENDS 210bp 3' OF EXON 11

```
gaattagacg aggcgatcag gttggtcaat gtatccaatc agcagtatgg ccagattctc      60
cagatgaccc ggaagcactt ggaggacacc gcctatctgg tggagaagat gagagggcaa     120
tttggctggg tgtctgaact ggcaaaccag gccccagaaa cagagatcat ctttaattca     180
atacaggtaa gaagatctaa tgcatcctat atccagtaag t                         221
```

FIG.21A

Met Lys Ile Lys Ala Glu Lys Asn Glu Gly Pro Ser Arg Ser Trp Trp
1 5 10 15
Gln Leu His Trp Gly Asp Ile Ala Asn Asn Ser Gly Asn Met Lys Pro
20 25 30
Pro Leu Leu Val Phe Ile Val Cys Leu Leu Trp Leu Lys Asp Ser His
35 40 45
Cys Ala Pro Thr Trp Lys Asp Lys Thr Ala Ile Ser Glu Asn Leu Lys
50 55 60
Ser Phe Ser Glu Val Gly Glu Ile Asp Ala Asp Glu Glu Val Lys Lys
65 70 75 80
Ala Leu Thr Gly Ile Lys Gln Met Lys Ile Met Met Glu Arg Lys Glu
85 90 95
Lys Glu His Thr Asn Leu Met Ser Thr Leu Lys Lys Cys Arg Glu Glu
100 105 110
Lys Gln Glu Ala Leu Lys Leu Leu Asn Glu Val Gln Glu His Leu Glu
115 120 125
Glu Glu Glu Arg Leu Cys Arg Glu Ser Leu Ala Asp Ser Trp Gly Glu
130 135 140
Cys Arg Ser Cys Leu Glu Asn Asn Cys Met Arg Ile Tyr Thr Thr Cys
145 150 155 160
Gln Pro Ser Trp Ser Ser Val Lys Asn Lys Ile Glu Arg Phe Phe Arg
165 170 175
Lys Ile Tyr Gln Phe Leu Phe Pro Phe His Glu Asp Asn Glu Lys Asp
180 185 190
Leu Pro Ile Ser Glu Lys Leu Ile Glu Glu Asp Ala Gln Leu Thr Gln
195 200 205
Met Glu Asp Val Phe Ser Gln Leu Thr Val Asp Val Asn Ser Leu Phe
210 215 220
Asn Arg Ser Phe Asn Val Phe Arg Gln Met Gln Gln Glu Phe Asp Gln
225 230 235 240
Thr Phe Gln Ser His Phe Ile Ser Asp Thr Asp Leu Thr Glu Pro Tyr
245 250 255
Phe Phe Pro Ala Phe Ser Lys Glu Pro Met Thr Lys Ala Asp Leu Glu
260 265 270
Gln Cys Trp Asp Ile Pro Asn Phe Phe Gln Leu Phe Cys Asn Phe Ser
275 280 285
Val Ser Ile Tyr Glu Ser Val Ser Glu Thr Ile Thr Lys Met Leu Lys
290 295 300
Ala Ile Glu Asp Leu Pro Lys Gln Asp Lys Ala Pro Asp His Gly Gly
305 310 315 320
Leu Ile Ser Lys Met Leu Pro Gly Gln Asp Arg Gly Leu Cys Gly Glu
325 330 335
Leu Asp Gln Asn Leu Ser Arg Cys Phe Lys Phe His Glu Lys Cys Gln
340 345 350
Lys Cys Gln Ala His Leu Ser Glu Asp Cys Pro Asp Val Pro Ala Leu
355 360 365
His Thr Glu Leu Asp Glu Ala Ile Arg Leu Val Asn Val Ser Asn Gln
370 375 380

FIG.21B-1

```
Gln Tyr Gly Gln Ile Leu Gln Met Thr Arg Lys His Leu Glu Asp Thr
385                 390                 395                 400
Ala Tyr Leu Val Glu Lys Met Arg Gly Gln Phe Gly Trp Val Ser Glu
                405                 410                 415
Leu Ala Asn Gln Ala Pro Glu Thr Glu Ile Ile Phe Arg Arg Ser Asn
            420                 425                 430
Ala Ser Tyr Ile Gln
        435
```

FIG.21B-2

```
acacagaatt agacgaggcg atcaggttgg tcaatgtatc caatcagcag tatggccaga      60
ttctccagat gacccggaag cacttggagg acaccgccta tctggtggag aagatgagag     120
ggcaatttgg ctgggtgtct gaactggcaa accaggcccc agaaacagag atcatcttta     180
attcaataca ggtagttcca aggattcatg aaggaaatat ttccaaacaa gatgaaacaa     240
tgatgacaga cttaagcatt ctgccttcct ctaatttcac actcaagatc cctcttgaag     300
aaagtgctga gagttctaac ttcattggct acgtagtggc aaaagctcta cagcatttta     360
aggaacattt taaaacctgg taagcagagt gcctggttag gaatgccttg ttgacaggaa     420
tagttaattc tcaaaaggga aaaacaaaac ttgtttcaaa atacctggaa aacatgttta     480
acctcattaa taaagacatg aaaacaaaca agatggcatt ttct                      524
```

FIG.22

```
gaattagacg aggcgatcag gttggtcaat gtatccaatc agcagtatgg ccagattctc      60
cagatgaccc ggaagcactt ggaggacacc gcctatctgg tggagaagat gagagggcaa     120
tttggctggg tgtctgaact ggcaaaccag gccccagaaa cagagatcat ctttaattca     180
atacaggtag ttccaaggat tcatgaagga aatatttcca aacaagatga aacaatgatg     240
acagacttaa gcattctgcc ttcctctaat ttcacactca agatccctct tgaagaaagt     300
gctgagagtt ctaacttcat tggctacgta gtggcaaaag ctctacagca ttttaaggaa     360
cattttaaaa cctgaaaaag atcctgaggc tcagtgtcca aggtccaatg aactactcag     420
gtcggaggtg gtagagcagc atgtggagcc agttctctct ccgactccat catcacactg     480
cacggcttcc tgttaagata tttgctcaaa aaatgcgaga tataaaaatc tgggtaagaa     540
gatctaatgc atcctatatc cagtaagt                                       568
```

FIG.23A

```
Met Lys Ile Lys Ala Glu Lys Asn Glu Gly Pro Ser Arg Ser Trp Trp
 1               5                  10                  15
Gln Leu His Trp Gly Asp Ile Ala Asn Asn Ser Gly Asn Met Lys Pro
            20                  25                  30
Pro Leu Leu Val Phe Ile Val Cys Leu Leu Trp Leu Lys Asp Ser His
        35                  40                  45
Cys Ala Pro Thr Trp Lys Asp Lys Thr Ala Ile Ser Glu Asn Leu Lys
    50                  55                  60
Ser Phe Ser Glu Val Gly Glu Ile Asp Ala Asp Glu Glu Val Lys Lys
65                  70                  75                  80
Ala Leu Thr Gly Ile Lys Gln Met Lys Ile Met Met Glu Arg Lys Glu
                85                  90                  95
Lys Glu His Thr Asn Leu Met Ser Thr Leu Lys Lys Cys Arg Glu Glu
            100                 105                 110
Lys Gln Glu Ala Leu Lys Leu Leu Asn Glu Val Gln Glu His Leu Glu
        115                 120                 125
Glu Glu Glu Arg Leu Cys Arg Glu Ser Leu Ala Asp Ser Trp Gly Glu
    130                 135                 140
Cys Arg Ser Cys Leu Glu Asn Asn Cys Met Arg Ile Tyr Thr Thr Cys
145                 150                 155                 160
Gln Pro Ser Trp Ser Ser Val Lys Asn Lys Ile Glu Arg Phe Phe Arg
                165                 170                 175
Lys Ile Tyr Gln Phe Leu Phe Pro Phe His Glu Asp Asn Glu Lys Asp
            180                 185                 190
Leu Pro Ile Ser Glu Lys Leu Ile Glu Glu Asp Ala Gln Leu Thr Gln
        195                 200                 205
Met Glu Asp Val Phe Ser Gln Leu Thr Val Asp Val Asn Ser Leu Phe
    210                 215                 220
Asn Arg Ser Phe Asn Val Phe Arg Gln Met Gln Gln Glu Phe Asp Gln
225                 230                 235                 240
Thr Phe Gln Ser His Phe Ile Ser Asp Thr Asp Leu Thr Glu Pro Tyr
                245                 250                 255
Phe Phe Pro Ala Phe Ser Lys Glu Pro Met Thr Lys Ala Asp Leu Glu
            260                 265                 270
Gln Cys Trp Asp Ile Pro Asn Phe Phe Gln Leu Phe Cys Asn Phe Ser
        275                 280                 285
Val Ser Ile Tyr Glu Ser Val Ser Glu Thr Ile Thr Lys Met Leu Lys
    290                 295                 300
Ala Ile Glu Asp Leu Pro Lys Gln Asp Lys Ala Pro Asp His Gly Gly
305                 310                 315                 320
Leu Ile Ser Lys Met Leu Pro Gly Gln Asp Arg Gly Leu Cys Gly Glu
                325                 330                 335
Leu Asp Gln Asn Leu Ser Arg Cys Phe Lys Phe His Glu Lys Cys Gln
            340                 345                 350
Lys Cys Gln Ala His Leu Ser Glu Asp Cys Pro Asp Val Pro Ala Leu
        355                 360                 365
His Thr Glu Leu Asp Glu Ala Ile Arg Leu Val Asn Val Ser Asn Gln
    370                 375                 380
Gln Tyr Gly Gln Ile Leu Gln Met Thr Arg Lys His Leu Glu Asp Thr
385                 390                 395                 400
```

FIG.23B

```
Ala Tyr Leu Val Glu Lys Met Arg Gly Gln Phe Gly Trp Val Ser Glu
                405                 410                 415
Leu Ala Asn Gln Ala Pro Glu Thr Glu Ile Ile Phe Asn Ser Ile Gln
            420                 425                 430
Val Val Pro Arg Ile His Glu Gly Asn Ile Ser Lys Gln Asp Glu Thr
        435                 440                 445
Met Met Thr Asp Leu Ser Ile Leu Pro Ser Ser Asn Phe Thr Leu Lys
    450                 455                 460
Ile Pro Leu Glu Glu Ser Ala Glu Ser Ser Asn Phe Ile Gly Tyr Val
465                 470                 475                 480
Val Ala Lys Ala Leu Gln His Phe Lys Glu His Phe Lys Thr
                485                 490
```

FIG.23C

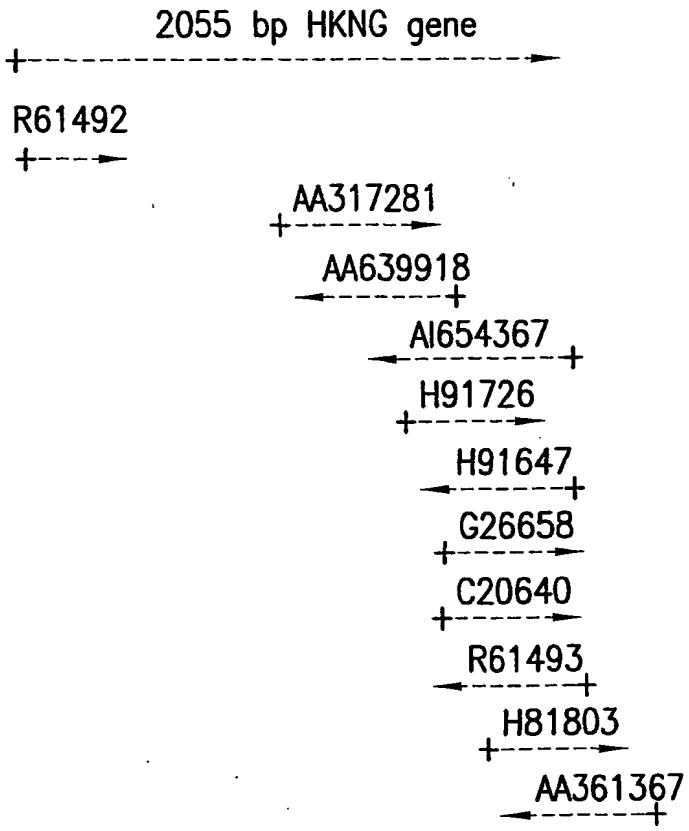

2055 bp HKNG gene
R61492
AA317281
AA639918
AI654367
H91726
H91647
G26658
C20640
R61493
H81803
AA361367

FIG.24

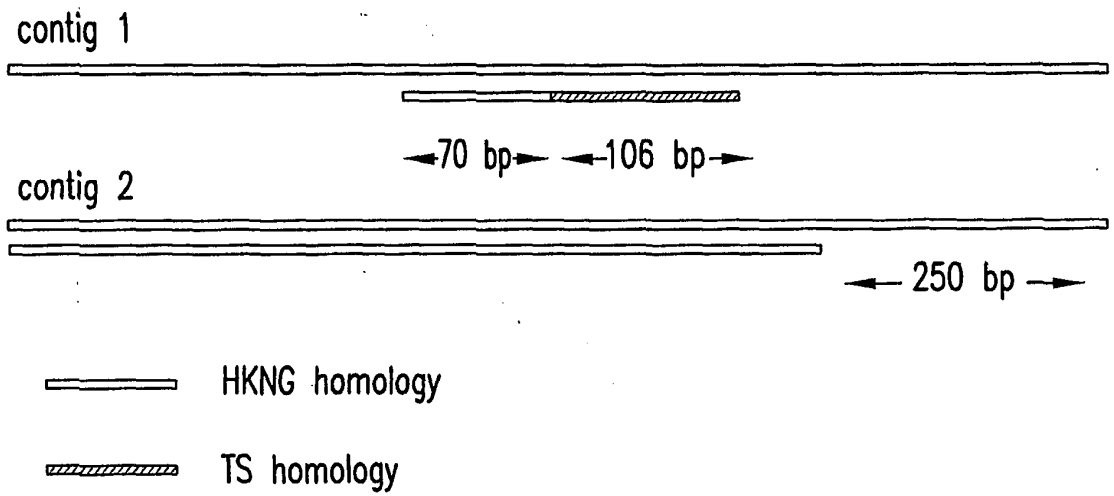

contig 1
70 bp
106 bp
contig 2
250 bp
HKNG homology
TS homology

FIG.25

GGTGTCTATG TTCTATCACA TCTACAAACA TGTCACTTCC TAATTAACAA AATGTTCTTC
CTTTAGTTTG CTTTTGCACT TAAAATATAT ATAATTGACT TTTTTGGAAA AAAATCTAAG
ATTCATTGCT TTGTTTTGTA AAGACCAATA GGTTCTGTAT AGTCTTTTTT TAAATTGTGG
TAAAATACAC ATGGCATTAA TTTACCATTT TAACCATTTT AAAGTGCACA ATTTGTGGCA
TTAAGTACAC TCACGTTGCT GTGCAACCAT CACCACCGTC CATCTTCAGA ACCTTTTTAT
CTTCCTAAAC TGAAACTCTG TACTCGTTAA GCACTCACTT CCCTTTTCCC CATCCCCCAG
CCCGTAGCAA CCACGACTGT ACTTTCTATG AATTTGACTA CTCTAGGTAC TGCATGTAGG
TGGAATCATA CAGTATTTGT CTTTTGCTTTG KTTTGKTTTG TTTTTTGTTT TCTAAGACAG
GGTCTCACTC TGTCGCCCTA GCTGGATTGC AGAGTTAAGT TTATGATTAT GAAATAAAAA
CTAAATAACN ATTGTCCTCG TTTG

FIG.26

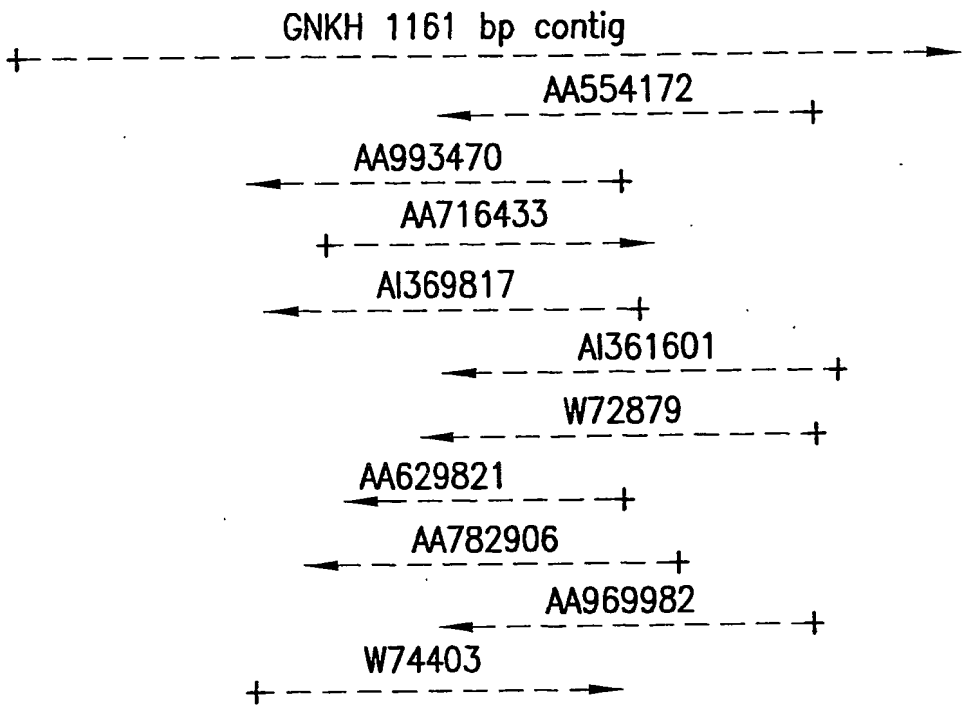

GNKH 1161 bp contig
AA554172
AA993470
AA716433
AI369817
AI361601
W72879
AA629821
AA782906
AA969982
W74403

FIG.27

```
cctgaaagcc tggcgccaat gacccgcgag acatttttg cctggggtgc tcctgtcgga      60
aaggaaagag gaaaggacga ctaagaactt atactcgaac tcccgaattt ctcttttcaa     120
ggtttaagag gaaagctggt tcgtggggat tggatgggag gccaccagga aaccaagttc     180
ccgcgccagc ttcagtgctc tcctcttycc gccgcctttg ccccgcccac atcactttcg     240
ctccagtttt tgaaaacgct gcgaagcgga atggtccaca ggggaaaacg gaggaggggc     300
caaagccagg actttgagac cggcgcgcgg tcaagcccag gcagctctcc ctaaccctcc     360
agcactgggc aaacgctgcc cgatgacgcc cgcctcgggg gccacggcat cactggggcg     420
actgcgagcc cggccgcgga gccgctggga cgcggcttac ctcccggctg tcgctgctgt     480
gtgtgttgcc cgcgccagtc acgtccctaa tgggaccctc cgtttcggcg tctgtaaggc     540
gaggaggacg atgcgtcccc tccctsgcag gattgaggtt aggactaaac ggggtccgca     600
gcgcccggca gctcccgagc gctctcccca gccgcgcctc cctccttccc gccacccgtc     660
ccgcaggggc ccgcggcgtc acctctcagg ctgtagcgcg cctgcatgcc gaataccgac     720
agggtgccgg tgcccgtgcg gtcgtccttc ctgacgccgc agcggaggat gtgttggatc     780
tgccccagga tttccaggtc ccagatgaag agataattct acttactgga tataggatgc     840
attagatctt cttaccttaa aaaaaaaaaa aaaggcagca atgatcaaaa tactaataaa     900
ttactcacag actcagtgta ttttttcttg gagtaaaagt ccaggatggg taatagaata     960
cctgctgttg gcttttggaa aaattggtac tgtatgtagc aaaataatgt gaaacccata    1020
tgcatggata ttcttaacaa tttgaagaaa tcgtcacagc tttcctgggt tgttgagcct    1080
ctaaaatggt cttttcctct gatgtgataa taaagtgttt attttgaact caaaaaaaaa    1140
aaaaaaaaaa aaaaaaaaaa a                                              1161
```

FIG.28

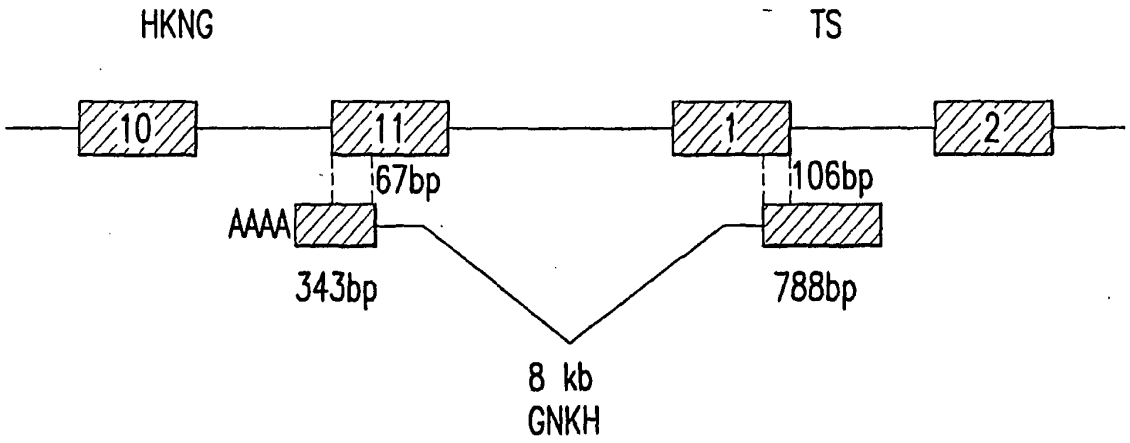


FIG.29

HKNG GENOMIC vs gnkhexp:

```
HKNG GENOMIC   888  cctgaaagcctggcgccaatgacccgcgagacatttttgcctggggtg  936
                    ||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp     1  cctgaaagcctggcgccaatgacccgcgagacatttttgcctggggtg   49

HKNG GENOMIC   937 ctcctgtcggaaaggaaagaggaaaggacgactaag-----a-actcgaa  980
                    |||||||||||||||||||||||||||||||||||||| |||||||
    gnkh exp    50 ctcctgtcggaaaggaaagaggaaaggacgactaagaacttatactcgaa   99

HKNG GENOMIC   981 ctcccgaatttctcttttcaaggtttaagaggaaagctggttcgtgggga 1030
                    ||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   100 ctcccgaatttctcttttcaaggtttaagaggaaagctggttcgtgggga  149

HKNG GENOMIC  1031 ttggatgggaggccaccaggaaaccaagttcccgcgccagcttcagtgct 1080
                    ||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   150 ttggatgggaggccaccaggaaaccaagttcccgcgccagcttcagtgct  199

HKNG GENOMIC  1081 stcctcttcccgccgcctttgccccgcccacatcactttcgctccagttt 1130
                    ||||||| ||||||||||||||||||||||||||||||||||||||||
    gnkh exp   200 ctcctcttyccgccgcctttgccccgcccacatcactttcgctccagttt  249

HKNG GENOMIC  1131 ttgaaaacgctgcgaagcggaatggtccacaggggaaaacggaggagggg 1180
                    ||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   250 ttgaaaacgctgcgaagcggaatggtccacaggggaaaacggaggagggg  299

HKNG GENOMIC  1181 ccaaagccaggactttgagaccggcgcgcggtcaagcccaggcagctctc 1230
                    ||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   300 ccaaagccaggactttgagaccggcgcgcggtcaagcccaggcagctctc  349

HKNG GENOMIC  1231 cctaaccctccagcactgggcaaacgctgcccgatgacgcccgcctcggg 1280
                    ||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   350 cctaaccctccagcactgggcaaacgctgcccgatgacgcccgcctcggg  399

HKNG GENOMIC  1281 ggccacggcatcactggggcgactgcgagcccggccgcggagccgctggg 1330
                    ||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   400 ggccacggcatcactggggcgactgcgagcccggccgcggagccgctggg  449

HKNG GENOMIC  1331 acgcggcttacctcccggctgtcgctgctgtgtgtgttgcccgcgccagt 1380
                    ||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   450 acgcggcttacctcccggctgtcgctgctgtgtgtgttgcccgcgccagt  499

HKNG GENOMIC  1381 cacgtccctaatgggaccctccgtttcggcgtctgtaaggcgaggaggac 1430
                    ||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   500 cacgtccctaatgggaccctccgtttcggcgtctgtaaggcgaggaggac  549

HKNG GENOMIC  1431 gatgcgtcccctccctggcaggattgaggttaggactaaacggggtccgc 1480
                    |||||||||||||||| ||||||||||||||||||||||||||||||
    gnkh exp   550 gatgcgtcccctccctsgcaggattgaggttaggactaaacggggtccgc  599

HKNG GENOMIC  1481 agcgcccggcagctcccgagcgctctccccagccgcgcctccctccttcc 1530
                    ||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   600 agcgcccggcagctcccgagcgctctccccagccgcgcctccctccttcc  649
```

FIG.30A

```
HKNG GENOMIC  1531 cgccacccgtcccgcaggggcccgcggcgtcacctctcaggctgtagcgc 1580
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   650 cgccacccgtcccgcaggggcccgcggcgtcacctctcaggctgtagcgc  699

HKNG GENOMIC  1581 gcctgcatgccgaataccgacagggtgccggtgcccgtgcggtcgtcctt 1630
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   700 gcctgcatgccgaataccgacagggtgccggtgcccgtgcggtcgtcctt  749

HKNG GENOMIC  1631 cctgacgccgcagcggaggatgtgttggatctgccccaggtact..... 1669
                   ||||||||||||||||||||||||||||||||||||>>>>> 7882
    gnkh exp   750 cctgacgccgcagcggaggatgtgttggatctgccccag.....  788

HKNG GENOMIC  1669 ttcaggatttccaggtcccagatgaagagataattctacttactggatat 9596
                   >>>>>||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   788 .....gatttccaggtcccagatgaagagataattctacttactggatat  833

HKNG GENOMIC  9597 aggatgcattagatcttcttaccttaaaaaaaaaaaaaaa-gcagcaatg 9645
                   |||||||||||||||||||||||||||||||||||||| |||||||||
    gnkh exp   834 aggatgcattagatcttcttaccttaaaaaaaaaaaaaaaggcagcaatg  883

HKNG GENOMIC  9646 atcaaaatactaataaattactcacagactcagtgtatttttcttggag 9695
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   884 atcaaaatactaataaattactcacagactcagtgtatttttcttggag  933

HKNG GENOMIC  9696 taaaagtccaggatgggtaatagaatacctgctgttggcttttggaaaaa 9745
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp   934 taaaagtccaggatgggtaatagaatacctgctgttggcttttggaaaaa  983

HKNG GENOMIC  9746 ttggtactgtgtgtagcaaaataatgtgaaacccatatgcatggatattc 9795
                   |||||||||| ||||||||||||||||||||||||||||||||||||||||
    gnkh exp   984 ttggtactgtatgtagcaaaataatgtgaaacccatatgcatggatattc 1033

HKNG GENOMIC  9796 ttaacaatttgaagaaatcgtcacagctttcctgggttgttgagcctcta 9845
                   ||||||||||||||||||||||||||||||||||||||||||||||||||
    gnkh exp  1034 ttaacaatttgaagaaatcgtcacagctttcctgggttgttgagcctcta 1083

HKNG GENOMIC  9846 agatggtcttttcctctgatgtgataataaagtgtttattctgaactc   9893
                   | ||||||||||||||||||||||||||||||||||||| |||||||
    gnkh exp  1084 aaatggtcttttcctctgatgtgataataaagtgtttattttgaactc   1131
```

FIG.30B

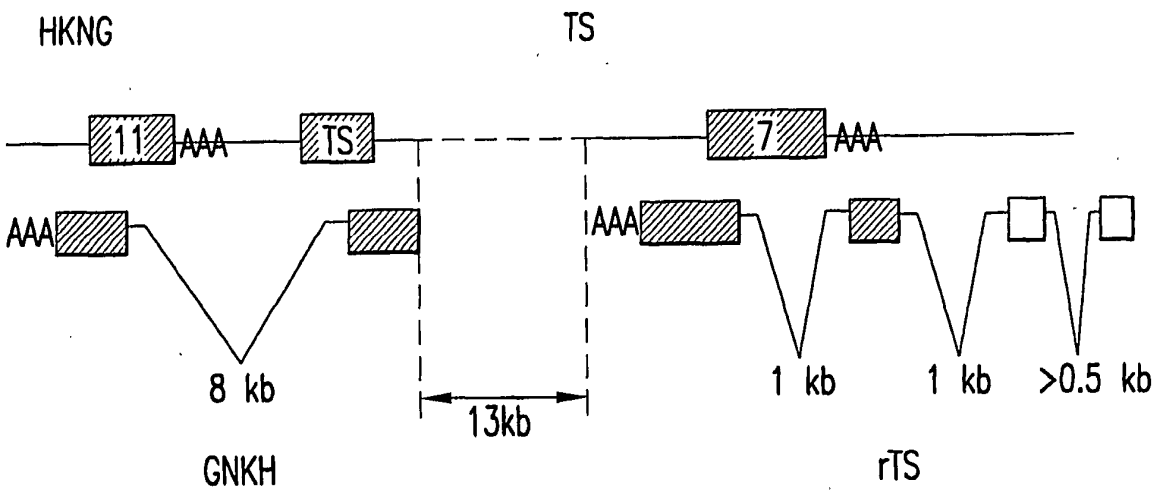


FIG.31

MTPASGATASLGRLRARPRSRWDAAYLPAVAAVCVARASHVPNGTLRFGVCKARRTMRPLPXRIEVRTKRGPQRPAAPER
SPQPRLPPSRHPSRRGPRRHLSGCSAPACRIPTGCRCPCGRPS

FIG.32

MGPSVSASVRRGGRCVPSLAGLRLGLNGVRSARQLPSALPSRASLLPATRPAGARGVTSQAVARLHAEYRQGAGARAVVL
PDAAAEDVLDLPQDFQVPDEEIILLTGYRMH

FIG.33

```
   S   Q   L   A   S   H   N   P   V   T   E   D   I   F   N   S   T   K   A      19
G TCT CAA CTG GCA AGC CAT AAC CCA GTG ACT GAG GAC ATC TTT AAT TCA ACA AAG GCA      58

 V   P   K   I   H   G   G   D   S   S   K   Q   D   E   I   M   V   D   S   S    39
GTT CCA AAG ATT CAT GGA GGA GAT TCT TCC AAG CAG GAT GAA ATT ATG GTA GAC TCA AGC  118
```

FIG.34

```
human        fgwvSELANQAPETEIIFNSIQVVPRI--HEGNISKQDETMMTDLS (i/l) pssnf
bovine       fgwvTELASQTPGSENIFSFIKVVPGV--HEGNFSKQDE-KMIDIS (i/l) pssnf
guinea pig   fgwvLELAYQSPGAEDIFNPVKVMVALSAHEGNSSDQDD-TVVPSS (i/l) pssnf
rat          fgwvSQLASHNPVTEDIFNSTKAVPKI--HGGDSSKQDE-IMVDSS (i/l) pssnf
```

FIG.35

```
cataacccag tgactgagga catctttaat tcaacaaagg cagttccaaa gattcatgga      60
ggagattctt ccaagcagga tgaaattatg gtagactcaa gcagcattct gccttcctct     120
aacttcaccg tccagaatcc tcctgaagaa ggtgctgaga gctcaaatgt tatttactac     180
atggcagcta aagttctgca gcatctaaag ggatgttttg aaacttggta agaatagctg     240
attaggaaag ctttgttgag agggtaggta acataaaaaa aaaaaaaaa                 289
```

FIG.36A

```
His Asn Pro Val Thr Glu Asp Ile Phe Asn Ser Thr Lys Ala Val Pro
 1               5                  10                  15
Lys Ile His Gly Gly Asp Ser Ser Lys Gln Asp Glu Ile Met Val Asp
            20                  25                  30
Ser Ser Ser Ile Leu Pro Ser Ser Asn Phe Thr Val Gln Asn Pro Pro
        35                  40                  45
Glu Glu Gly Ala Glu Ser Ser Asn Val Ile Tyr Tyr Met Ala Ala Lys
    50                  55                  60
Val Leu Gln His Leu Lys Gly Cys Phe Glu Thr Trp Glu Leu Ile Arg
65                  70                  75                  80
Lys Ala Leu Leu Arg Gly Val Thr Lys Lys Lys Lys
                85                  90
```

FIG.36B

```
cccttcactg cgcgcccact gggaaggaga cagatgctac ggatggaaac ctaaagagtc      60
ttccagaggt aggagaggca gatgtagagg gagaggtcaa gaaggctttg attggcatta     120
agcaaatgaa aatcatgatg gaaaggagag aggaggaaca cgcaaaattg atgaaagcct     180
tgaagaagtg caaagaagaa aagcaggagg cccagaaact catgaacgaa gtgcaagaac     240
gtctggagga agaagaaaag ctatgtcagg catcttctat aggttcttgg gatggatgca     300
ggccatgttt ggaaagtaac tgcatacgat tttatacagc ttgccaacct ggttggtcct     360
ctgtgaaaag catgatgaag caatttctca agaagatata ccgatttctg tcttcccaga     420
gtgaagatgt aaaggatccc cctgccatag aacagctgac taaggaagat ttacaagtgg     480
tacacataga gaacctgttt agccagctgg ccgtggatgc aaaatctctc ttcaacatga     540
gcttttacat ttttaagcag atgcagcaag aatttgatca ggcttttcaa ttatacttca     600
tgtccgatgt ggacttaatg gagccatacc ccccagcttt atctaaagag ataatcaaaa     660
aagaagaact tgggcaaagg tggggcattc ccaatgtctt ccagctgttt cataatttca     720
gtctctctgt ttatgggaga gtccaacaaa taataatgaa gacactcaat gcaattgaag     780
attcatggga accacacaaa gagttagacc agagaggtat gacttcagag atgttacctg     840
agcaaaatgg agaaatgtgt gaggaatttg tcaagaattt atctggatgt ttaaaatttc     900
gtaaaagatg ccaaaaatgt cacaattacc tatctgaaga atgccctgat gtacctgaac     960
ttcacataga attccttgag gccctgaaat tagtcaatgt atccaatcag caatatgatc    1020
agattgtcca gatgacccag tatcatttgg aagataccat atacctgatg gagaaaatgc    1080
aagagcagtt tggatgggtg tctcaactgg caagccataa                          1120
```

FIG.37A

Leu His Cys Ala Pro Thr Gly Lys Glu Thr Asp Ala Thr Asp Gly Asn
1 5 10 15
Leu Lys Ser Leu Pro Glu Val Gly Glu Ala Asp Val Glu Gly Glu Val
20 25 30
Lys Lys Ala Leu Ile Gly Ile Lys Gln Met Lys Ile Met Met Glu Arg
35 40 45
Arg Glu Glu Glu His Ala Lys Leu Met Lys Ala Leu Lys Lys Cys Lys
50 55 60
Glu Glu Lys Gln Glu Ala Gln Lys Leu Met Asn Glu Val Gln Glu Arg
65 70 75 80
Leu Glu Glu Glu Glu Lys Leu Cys Gln Ala Ser Ser Ile Gly Ser Trp
85 90 95
Asp Gly Cys Arg Pro Cys Leu Glu Ser Asn Cys Ile Arg Phe Tyr Thr
100 105 110
Ala Cys Gln Pro Gly Trp Ser Ser Val Lys Ser Met Met Lys Gln Phe
115 120 125
Leu Lys Lys Ile Tyr Arg Phe Leu Ser Ser Gln Ser Glu Asp Val Lys
130 135 140
Asp Pro Pro Ala Ile Glu Gln Leu Thr Lys Glu Asp Leu Gln Val Val
145 150 155 160
His Ile Glu Asn Leu Phe Ser Gln Leu Ala Val Asp Ala Lys Ser Leu
165 170 175
Phe Asn Met Ser Phe Tyr Ile Phe Lys Gln Met Gln Gln Glu Phe Asp
180 185 190
Gln Ala Phe Gln Leu Tyr Phe Met Ser Asp Val Asp Leu Met Glu Pro
195 200 205
Tyr Pro Pro Ala Leu Ser Lys Glu Ile Ile Lys Lys Glu Glu Leu Gly
210 215 220
Gln Arg Trp Gly Ile Pro Asn Val Phe Gln Leu Phe His Asn Phe Ser
225 230 235 240
Leu Ser Val Tyr Gly Arg Val Gln Gln Ile Ile Met Lys Thr Leu Asn
245 250 255
Ala Ile Glu Asp Ser Trp Glu Pro His Lys Glu Leu Asp Gln Arg Gly
260 265 270
Met Thr Ser Glu Met Leu Pro Glu Gln Asn Gly Glu Met Cys Glu Glu
275 280 285
Phe Val Lys Asn Leu Ser Gly Cys Leu Lys Phe Arg Lys Arg Cys Gln
290 295 300
Lys Cys His Asn Tyr Leu Ser Glu Glu Cys Pro Asp Val Pro Glu Leu
305 310 315 320
His Ile Glu Phe Leu Glu Ala Leu Lys Leu Val Asn Val Ser Asn Gln
325 330 335
Gln Tyr Asp Gln Ile Val Gln Met Thr Gln Tyr His Leu Glu Asp Thr
340 345 350
Ile Tyr Leu Met Glu Lys Met Gln Glu Gln Phe Gly Trp Val Ser Gln
355 360 365
Leu Ala Ser His Asn Pro Val Thr Glu Asp Ile Phe Asn Ser Thr Lys
370 375 380
Ala Val Pro Lys Ile His Gly Gly Asp Ser Ser Lys Gln
385 390 395

FIG.37B

```
ttttttttttt ttttttttcaa ggctttcatc aattttgcgt gttcctcctc tctcctttcc      60
atcatgattt tcatttgctt aatgccaatc aaagccttct tgacctctcc ctctacatct     120
gcctctccta cctctggaag actctttagg tttccatccg tagcatctgt ctccttccaa     180
gtaggtgcac tgtcacaata tttcaaccat aacagataca cagaaatcac aaagagtggt     240
ggctgcatgg tccagtgttc caccgatatt gcagctctcc ccagagaaat tgccactaac     300
ttctgaaagg accttcactt tttacgatgt gcctcgtgcc g                         341
```

FIG.38A

```
cggcacgagg cacatcgtaa aaagtgaagg tcctttcaga agttagtggc aatttctctg      60
gggagagctg caatatcggt ggaacactgg accatgcagc caccactctt tgtgatttct     120
gtgtatctgt tatggttgaa atattgtgac agtgcaccta cttggaagga gacagatgct     180
acggatggaa acctaaagag tcttccagag gtaggagagg cagatgtaga gggagaggtc     240
aagaaggctt tgattggcat taagcaaatg aaaatcatga tggaaaggag agaggaggaa     300
cacgcaaaat tgatgaaagc cttgaaaaaa aaaaaaaaaa a                         341
```

FIG.38B

```
Arg His Glu Ala His Arg Lys Lys Arg Ser Phe Gln Lys Leu Val Ala
 1               5                  10                  15
Ile Ser Leu Gly Arg Ala Ala Ile Ser Val Glu His Trp Thr Met Gln
            20                  25                  30
Pro Pro Leu Phe Val Ile Ser Val Tyr Leu Leu Trp Leu Lys Tyr Cys
        35                  40                  45
Asp Ser Ala Pro Thr Trp Lys Glu Thr Asp Ala Thr Asp Gly Asn Leu
    50                  55                  60
Lys Ser Leu Pro Glu Val Gly Glu Ala Asp Val Glu Gly Glu Val Lys
65                  70                  75                  80
Lys Ala Leu Ile Gly Ile Lys Gln Met Lys Ile Met Met Glu Arg Arg
                85                  90                  95
Glu Glu Glu His Ala Lys Leu Met Lys Ala Leu Lys Lys Lys Lys Lys
            100                 105                 110
```

FIG.38C

```
ggcaccgagg cacatcgtaa aaagtgaagg tcctttcaga agttagtggc aatttctctg      60
gggagagctg caatatcggt ggaacactgg accatgcagc caccactctt tgtgatttct     120
gtgtatctgt tatggtgaaa tattgtgaca gtgcacctac ttggaaggag acagatgcta     180
cggatggaaa cctaaagagt cttccagagg taggagaggc agatgtagag ggagaggtca     240
agaaggcttt gattggcatt aagcaaatga aaatcatgat ggaaaggaga gaggaggaac     300
acgcaaaatt gatgaaagcc ttgaagaagt gcaaagaaga aaagcaggag gcccagaaac     360
tcatgaacga agtgcaagaa cgtctggagg aagaagaaaa gctatgtcag gcatcttcta     420
taggttcttg ggatggatgc aggccatgtt tggaaagtaa ctgcatacga ttttatacag     480
cttgccaacc tggttggtcc tctgtgaaaa gcatgatgaa gcaatttctc aagaagatat     540
accgatttct gtcttcccag agtgaagatg taaaggatcc ccctgccata gaacagctga     600
ctaaggaaga tttacaagtg gtacacatag agaacctgtt tagccagctg gccgtggatg     660
caaaatctct cttcaacatg agcttttaca tttttaagca gatgcagcaa gaatttgatc     720
aggcttttca attatacttc atgtccgatg tggacttaat ggagccatac cccccagctt     780
tatctaaaga gataatcaaa aaagaagaac ttgggcaaag gtggggcatt cccaatgtct     840
tccagctgtt tcataatttc agtctctctg tttatgggag agtccaacaa ataataatga     900
agacactcaa tgcaattgaa gattcatggg aaccacacaa agagttagac cagagaggta     960
tgacttcaga gatgttacct gagcaaaatg gagaaatgtg tgaggaattt gtcaagaatt    1020
tatctggatg tttaaaattt cgtaaaagat gccaaaaatg tcacaattac ctatctgaag    1080
aatgccctga tgtacctgaa cttcacatag aattccttga ggccctgaaa ttagtcaatg    1140
tatccaatca gcaatatgat cagattgtcc agatgaccca gtatcatttg gaagatacca    1200
tatacctgat ggagaaaatg caagagcagt ttggatgggt gtctcaactg gcaagccata    1260
acccagtgac tgaggacatc tttaattcaa caaaggcagt tccaaagatt catggaggag    1320
attcttccaa gcaggatgaa attatggtag actcaagcag cattctgcct tcctctaact    1380
tcaccgtcca gaatcctcct gaagaaggtg ctgagagctc aaatgttatt tactacatgg    1440
cagctaaagt tctgcagcat ctaaagggat gttttgaaac ttggtaagaa tagctgatta    1500
ggaaagcttt gttgagaggg taggtaacat aaaaaaaaaa aaaaa                    1545
```

FIG.39A

```
His Arg Gly Thr Ser Glx Lys Val Lys Val Leu Ser Glu Val Ser Gly
 1               5                  10                  15
Asn Phe Ser Gly Glu Ser Cys Asn Ile Gly Gly Thr Leu Asp His Ala
            20                  25                  30
Ala Thr Thr Leu Cys Asp Phe Cys Val Ser Val Met Val Lys Tyr Cys
        35                  40                  45
Asp Ser Ala Pro Thr Trp Lys Glu Thr Asp Ala Thr Asp Gly Asn Leu
    50                  55                  60
Lys Ser Leu Pro Glu Val Gly Glu Ala Asp Val Glu Gly Glu Val Lys
65                  70                  75                  80
Lys Ala Leu Ile Gly Ile Lys Gln Met Lys Ile Met Met Glu Arg Arg
                85                  90                  95
Glu Glu Glu His Ala Lys Leu Met Lys Ala Leu Lys Lys Cys Lys Glu
            100                 105                 110
Glu Lys Gln Glu Ala Gln Lys Leu Met Asn Glu Val Gln Glu Arg Leu
        115                 120                 125
Glu Glu Glu Glu Lys Leu Cys Gln Ala Ser Ser Ile Gly Ser Trp Asp
    130                 135                 140
Gly Cys Arg Pro Cys Leu Glu Ser Asn Cys Ile Arg Phe Tyr Thr Ala
145                 150                 155                 160
Cys Gln Pro Gly Trp Ser Ser Val Lys Ser Met Met Lys Gln Phe Leu
                165                 170                 175
Lys Lys Ile Tyr Arg Phe Leu Ser Ser Gln Ser Glu Asp Val Lys Asp
            180                 185                 190
Pro Pro Ala Ile Glu Gln Leu Thr Lys Glu Asp Leu Gln Val Val His
        195                 200                 205
Ile Glu Asn Leu Phe Ser Gln Leu Ala Val Asp Ala Lys Ser Leu Phe
    210                 215                 220
Asn Met Ser Phe Tyr Ile Phe Lys Gln Met Gln Gln Glu Phe Asp Gln
225                 230                 235                 240
Ala Phe Gln Leu Tyr Phe Met Ser Asp Val Asp Leu Met Glu Pro Tyr
                245                 250                 255
Pro Pro Ala Leu Ser Lys Glu Ile Ile Lys Lys Glu Glu Leu Gly Gln
            260                 265                 270
Arg Trp Gly Ile Pro Asn Val Phe Gln Leu Phe His Asn Phe Ser Leu
        275                 280                 285
Ser Val Tyr Gly Arg Val Gln Gln Ile Ile Met Lys Thr Leu Asn Ala
    290                 295                 300
Ile Glu Asp Ser Trp Glu Pro His Lys Glu Leu Asp Gln Arg Gly Met
305                 310                 315                 320
Thr Ser Glu Met Leu Pro Glu Gln Asn Gly Glu Met Cys Glu Glu Phe
                325                 330                 335
Val Lys Asn Leu Ser Gly Cys Leu Lys Phe Arg Lys Arg Cys Gln Lys
            340                 345                 350
Cys His Asn Tyr Leu Ser Glu Glu Cys Pro Asp Val Pro Glu Leu His
        355                 360                 365
Ile Glu Phe Leu Glu Ala Leu Lys Leu Val Asn Val Ser Asn Gln Gln
    370                 375                 380
Tyr Asp Gln Ile Val Gln Met Thr Gln Tyr His Leu Glu Asp Thr Ile
385                 390                 395                 400
```

FIG.39B-1

```
Tyr Leu Met Glu Lys Met Gln Glu Gln Phe Gly Trp Val Ser Gln Leu
                405                 410                 415
Ala Ser His Asn Pro Val Thr Glu Asp Ile Phe Asn Ser Thr Lys Ala
            420                 425                 430
Val Pro Lys Ile His Gly Gly Asp Ser Ser Lys Gln Asp Glu Ile Met
        435                 440                 445
Val Asp Ser Ser Ser Ile Leu Pro Ser Ser Asn Phe Thr Val Gln Asn
    450                 455                 460
Pro Pro Glu Glu Gly Ala Glu Ser Ser Asn Val Ile Tyr Tyr Met Ala
465                 470                 475                 480
Ala Lys Val Leu Gln His Leu Lys Gly Cys Phe Glu Thr Trp Glu Leu
                485                 490                 495
Ile Arg Lys Ala Leu Leu Arg Gly Asn Val Thr Asn Lys Lys Lys Lys
            500                 505                 510
```

FIG.39B-2

```
aaaacgacgg ccagtgcggc acgaggcaca tcgtaaaaag tgaaggtcct ttcagaagtt      60
agtggcaatt tctctgggga gagctgcaat atcggtggaa cactggacca tgcagccacc     120
actctttgtg atttctgtgt atctgttatg gttgaaatat tgtgacagtg cacctacttg     180
gaaggagaca gatgctacgg atggaaacct aaagagtctt ccagaggtag gagaggcaga     240
tgtagaggga gaggtcaaga aggctttgat tggcattaag caaatgaaaa tcatgatgga     300
aaggagagag gaggaacacg caaaattgat gaaagccttg aagaagtgca aagaagaaaa     360
gcaggaggcc cagaaactca tgaacgaagt gcaagaacgt ctggaggaag aagaaaagct     420
atgtcaggca tcttctatag gttcttggga tggatgcagg ccatgtttgg aaagtaactg     480
catacgattt tatacagctt gccaacctgg ttggtcctct gtgaaaagca tgatgaagca     540
atttctcaag aagatatacc gatttctgtc ttcccagagt gaagatgtaa aggatccccc     600
tgccatagaa cagctgacta aggaagattt acaagtggta cacatagaga acctgtttag     660
ccagctggcc gtggatgcaa aatctctctt caacatgagc ttttacattt ttaagcagat     720
gcagcaagaa tttgatcagg cttttcaatt atacttcatg tccgatgtgg acttaatgga     780
gccatacccc ccagctttat ctaaagagat aatcaaaaaa gaagaacttg ggcaaaggtg     840
gggcattccc aatgtcttcc agctgtttca taatttcagt ctctctgttt atgggagagt     900
ccaacaaata ataatgaaga cactcaatgc aattgaagat tcatgggaac cacacaaaga     960
gttagaccag agaggtatga cttcagagat gttacctgag caaaatggag aaatgtgtga    1020
ggaatttgtc aagaatttat ctggatgttt aaaatttcgt aaaagatgcc aaaaatgtca    1080
caattaccta tctgaagaat gccctgatgt acctgaactt cacatagaat tccttgaggc    1140
cctgaaatta gtcaatgtat ccaatcagca atatgatcag attgtccaga tgacccagta    1200
tcatttggaa gataccatat acctgatgga gaaaatgcaa gagcagtttg gatgggtgtc    1260
tcaactggca agccataacc cagtgactga ggacatcttt aattcaacaa aggcagttcc    1320
aaagattcat ggaggagatt cttccaagca ggatgaaatt atggtagact caagcagcat    1380
tctgccttcc tctaacttca ccgtccagaa tcctcctgaa gaaggtgctg agagctcaaa    1440
tgttatttac tacatggcag ctaaagttct gcagcatcta aagggatgtt ttgaaacttg    1500
gtaagaatag ctgattagga aagctttgtt gagagggtag g                        1541
```

FIG.40A

Met Gln Pro Pro Leu Phe Val Ile Ser Val Tyr Leu Leu Trp Leu Lys
1 5 10 15
Tyr Cys Asp Ser Ala Pro Thr Trp Lys Glu Thr Asp Ala Thr Asp Gly
20 25 30
Asn Leu Lys Ser Leu Pro Glu Val Gly Glu Ala Asp Val Glu Gly Glu
35 40 45
Val Lys Lys Ala Leu Ile Gly Ile Lys Gln Met Lys Ile Met Met Glu
50 55 60
Arg Arg Glu Glu Glu His Ala Lys Leu Met Lys Ala Leu Lys Lys Cys
65 70 75 80
Lys Glu Glu Lys Gln Glu Ala Gln Lys Leu Met Asn Glu Val Gln Glu
85 90 95
Arg Leu Glu Glu Glu Glu Lys Leu Cys Gln Ala Ser Ser Ile Gly Ser
100 105 110
Trp Asp Gly Cys Arg Pro Cys Leu Glu Ser Asn Cys Ile Arg Phe Tyr
115 120 125
Thr Ala Cys Gln Pro Gly Trp Ser Ser Val Lys Ser Met Met Lys Gln
130 135 140
Phe Leu Lys Lys Ile Tyr Arg Phe Leu Ser Ser Gln Ser Glu Asp Val
145 150 155 160
Lys Asp Pro Pro Ala Ile Glu Gln Leu Thr Lys Glu Asp Leu Gln Val
165 170 175
Val His Ile Glu Asn Leu Phe Ser Gln Leu Ala Val Asp Ala Lys Ser
180 185 190
Leu Phe Asn Met Ser Phe Tyr Ile Phe Lys Gln Met Gln Gln Glu Phe
195 200 205
Asp Gln Ala Phe Gln Leu Tyr Phe Met Ser Asp Val Asp Leu Met Glu
210 215 220
Pro Tyr Pro Pro Ala Leu Ser Lys Glu Ile Ile Lys Lys Glu Glu Leu
225 230 235 240
Gly Gln Arg Trp Gly Ile Pro Asn Val Phe Gln Leu Phe His Asn Phe
245 250 255
Ser Leu Ser Val Tyr Gly Arg Val Gln Gln Ile Ile Met Lys Thr Leu
260 265 270
Asn Ala Ile Glu Asp Ser Trp Glu Pro His Lys Glu Leu Asp Gln Arg
275 280 285
Gly Met Thr Ser Glu Met Leu Pro Glu Gln Asn Gly Glu Met Cys Glu
290 295 300
Glu Phe Val Lys Asn Leu Ser Gly Cys Leu Lys Phe Arg Lys Arg Cys
305 310 315 320
Gln Lys Cys His Asn Tyr Leu Ser Glu Glu Cys Pro Asp Val Pro Glu
325 330 335
Leu His Ile Glu Phe Leu Glu Ala Leu Lys Leu Val Asn Val Ser Asn
340 345 350
Gln Gln Tyr Asp Gln Ile Val Gln Met Thr Gln Tyr His Leu Glu Asp
355 360 365
Thr Ile Tyr Leu Met Glu Lys Met Gln Glu Gln Phe Gly Trp Val Ser
370 375 380
Gln Leu Ala Ser His Asn Pro Val Thr Glu Asp Ile Phe Asn Ser Thr
385 390 395 400

FIG.40B-1

```
Lys Ala Val Pro Lys Ile His Gly Gly Asp Ser Ser Lys Gln Asp Glu
                405                 410                 415
Ile Met Val Asp Ser Ser Ser Ile Leu Pro Ser Ser Asn Phe Thr Val
            420                 425                 430
Gln Asn Pro Pro Glu Glu Gly Ala Glu Ser Ser Asn Val Ile Tyr Tyr
        435                 440                 445
Met Ala Ala Lys Val Leu Gln His Leu Lys Gly Cys Phe Glu Thr Trp
    450                 455                 460
```

FIG.40B-2

```
aaaacgacgg ccagtgcggc acgaggcaca tcgtaaaaag tgaaggtcct ttcagaagtt      60
agtggcaatt tctctgggga gagctgcaat atcggtggaa cactggacca tgcagccacc     120
actctttgtg atttctgtgt atctgttatg gttgaaatat tgtgacagtg cacctacttg     180
gaaggagaca gatgctacgg atggaaacct aaagagtctt ccagaggtag gagaggcaga     240
tgtagaggga gaggtcaaga aggctttgat tggcattaag caaatgaaaa tcatgatgga     300
aaggagagag gaggaacacg caaaattgat gaaagccttg aagaagtgca aagaagaaaa     360
gcaggaggcc cagaaactca tgaacgaagt gcaagaacgt ctggaggaag aagaaaagct     420
atgtcaggca tcttctatag gttcttggga tggatgcagg ccatgtttgg aaagtaactg     480
catacgattt tatacagctt gccaacctgg ttggtcctct gtgaaaagca tgatgaagca     540
atttctcaag aagatatacc gatttctgtc ttcccagagt gaagatgtaa aggatccccc     600
tgccatagaa cagctgacta aggaagattt acaagtggta cacatagaga acctgtttag     660
ccagctggcc gtggatgcaa aatctctctt caacatgagc ttttacattt ttaagcagat     720
gcagcaagaa tttgatcagg cttttcaatt atacttcatg tccgatgtgg acttaatgga     780
gccatacccc ccagctttat ctaaagagat aatcaaaaaa gaagaacttg ggcaaaggtg     840
gggcattccc aatgtcttcc agctgtttca taatttcagt ctctctgttt atgggagagt     900
ccaacaaata ataatgaaga cactcaatgc aattgaagat tcatgggaac cacacaaaga     960
gttagaccag agaggtatga cttcagagat gttacctgag caaaatggag aaatgtgtga    1020
ggaatttgtc aagaatttat ctggatgttt aaaatttcgt aaaagatgcc aaaaatgtca    1080
caattaccta tctgaagaat gccctgatgt acctgaactt cacatagaat tccttgaggc    1140
cctgaaatta gtcaatgtat ccaatcagca atatgatcag attgtccaga tgacccagta    1200
tcatttggaa gataccatat acctgatgga gaaaatgcaa gagcagtttg gatgggtgtc    1260
tcaactggca agccataacc cagtgactga ggacatcttt aattcaacaa aggcagttcc    1320
aaagattcat ggaggagatt cttccaagca ggatgaaatt atggtagact caagcagcat    1380
tctgccttcc tctaacttca ccgtccagaa tcctcctgaa gaaggtgctg agagctcaaa    1440
tgttatttac tacatggcag ctaaagttct gcagcatcta aagggatgtt ttgaaacttg    1500
gtaagaatag ctgattagga aagctttgtt gagagggtag g                        1541
```

FIG.41A

Met Gln Pro Pro Leu Phe Val Ile Ser Val Tyr Leu Leu Trp Leu Lys
1 5 10 15
Tyr Cys Asp Ser Ala Pro Thr Trp Lys Glu Thr Asp Ala Thr Asp Gly
20 25 30
Asn Leu Lys Ser Leu Pro Glu Val Gly Glu Ala Asp Val Glu Gly Glu
35 40 45
Val Lys Lys Ala Leu Ile Gly Ile Lys Gln Met Lys Ile Met Met Glu
50 55 60
Arg Arg Glu Glu Glu His Ala Lys Leu Met Lys Ala Leu Lys Lys Cys
65 70 75 80
Lys Glu Glu Lys Gln Glu Ala Gln Lys Leu Met Asn Glu Val Gln Glu
85 90 95
Arg Leu Glu Glu Glu Glu Lys Leu Cys Gln Ala Ser Ser Ile Gly Ser
100 105 110
Trp Asp Gly Cys Arg Pro Cys Leu Glu Ser Asn Cys Ile Arg Phe Tyr
115 120 125
Thr Ala Cys Gln Pro Gly Trp Ser Ser Val Lys Ser Met Met Lys Gln
130 135 140
Phe Leu Lys Lys Ile Tyr Arg Phe Leu Ser Ser Gln Ser Glu Asp Val
145 150 155 160
Lys Asp Pro Pro Ala Ile Glu Gln Leu Thr Lys Glu Asp Leu Gln Val
165 170 175
Val His Ile Glu Asn Leu Phe Ser Gln Leu Ala Val Asp Ala Lys Ser
180 185 190
Leu Phe Asn Met Ser Phe Tyr Ile Phe Lys Gln Met Gln Gln Glu Phe
195 200 205
Asp Gln Ala Phe Gln Leu Tyr Phe Met Ser Asp Val Asp Leu Met Glu
210 215 220
Pro Tyr Pro Pro Ala Leu Ser Lys Glu Ile Thr Lys Lys Glu Glu Leu
225 230 235 240
Gly Gln Arg Trp Gly Ile Pro Asn Val Phe Gln Leu Phe His Asn Phe
245 250 255
Ser Leu Ser Val Tyr Gly Arg Val Gln Gln Ile Ile Met Lys Thr Leu
260 265 270
Asn Ala Ile Glu Asp Ser Trp Glu Pro His Lys Glu Leu Asp Gln Arg
275 280 285
Gly Met Thr Ser Glu Met Leu Pro Glu Gln Asn Gly Glu Met Cys Glu
290 295 300
Glu Phe Val Lys Asn Leu Ser Gly Cys Leu Lys Phe Arg Lys Arg Cys
305 310 315 320
Gln Lys Cys His Asn Tyr Leu Ser Glu Glu Cys Pro Asp Val Pro Glu
325 330 335
Leu His Ile Glu Phe Leu Glu Ala Leu Lys Leu Val Asn Val Ser Asn
340 345 350
Gln Gln Tyr Asp Gln Ile Val Gln Met Thr Gln Tyr His Leu Glu Asp
355 360 365
Thr Ile Tyr Leu Met Glu Lys Met Gln Glu Gln Phe Gly Trp Val Ser
370 375 380
Gln Leu Ala Ser His Asn Pro Val Thr Glu Asp Ile Phe Asn Ser Thr
385 390 395 400

FIG.41B-1

Lys Ala Val Pro Lys Ile His Gly Gly Asp Ser Ser Lys Gln Asp Glu
405 410 415
Ile Met Val Asp Ser Ser Ser Ile Leu Pro Ser Ser Asn Phe Thr Val
420 425 430
Gln Asn Pro Pro Glu Glu Gly Ala Glu Ser Ser Asn Val Ile Tyr Tyr
435 440 445
Met Ala Ala Lys Val Leu Gln His Leu Lys Gly Cys Phe Glu Thr Trp
450 455 460

FIG.41B-2

```
aaaacgacgg ccagtgcggc acgaggcaca tcgtaaaaag tgaaggtcct ttcagaagtt      60
agtggcaatt tctctgggga gagctgcaat atcggtggaa cactggacca tgcagccacc     120
actctttgtg atttctgtgt atctgttatg gttgaaatat tgtgacagtg cacctacttg     180
gaaggagaca gatgctacgg atggaaacct aaagagtctt ccagaggtag gagaggcaga     240
tgtagaggga gaggtcaaga aggctttgat tggcattaag caaatgaaaa tcatgatgga     300
aaggagagag gaggaacacg caaaattgat gaaagccttg aagaagtgca aagaagaaaa     360
gcaggaggcc cagaaactca tgaacgaagt gcaagaacgt ctggaggaag aagaaaagct     420
atgtcaggca tcttctatag gttcttggga tggatgcagg ccatgtttgg aaagtaactg     480
catacgattt tatacagctt gccaacctgg ttggtcctct gtgaaaagca tgatgaagca     540
atttctcaag aagatatacc gatttctgtc ttcccagagt gaagatgtaa aggatccccc     600
tgccatagaa cagctgacta aggaagattt acaagtggta cacatagaga acctgtttag     660
ccagctggcc gtggatgcaa aatctctctt caacatgagc ttttacattt ttaagcagat     720
gcagcaagaa tttgatcagg cttttcaatt atacttcatg tccgatgtgg acttaatgga     780
gccatacccc ccagctttat ctaaagagat aaccaaaaaa gaagaacttg ggcaaaggtg     840
gggcattccc aatgtcttcc agctgtttca taatttcagt ctctctgttt atgggagagt     900
ccaacaaata ataatgaaga cactcaatgc aattgaagat tcatgggaac cacacaaaga     960
gttagaccag agaggtatga cttcagagat gttacctgag caaaatggag aaatgtgtga    1020
ggaatttgtc aagaatttat ctggatgttt aaaatttcgt aaaagatgcc aaaaatgtca    1080
caattaccta tctgaaggca gttccaaaga ttcatggagg agattcttcc aagcaggatg    1140
aaattatggt agactcaagc agcattctgc cttcctctaa cttcaccgtc cagaatcctc    1200
ctgaagaagg tgctgagagc tcaaatgtta tttactacat ggcagctaaa gttctgcagc    1260
atctaaaggg atgttttgaa acttggtaag aatagctgat taggaaagct ttgttgagag    1320
ggtagg                                                               1326
```

FIG.42A

Met Gln Pro Pro Leu Phe Val Ile Ser Val Tyr Leu Leu Trp Leu Lys
1 5 10 15
Tyr Cys Asp Ser Ala Pro Thr Trp Lys Glu Thr Asp Ala Thr Asp Gly
20 25 30
Asn Leu Lys Ser Leu Pro Glu Val Gly Glu Ala Asp Val Glu Gly Glu
35 40 45
Val Lys Lys Ala Leu Ile Gly Ile Lys Gln Met Lys Ile Met Met Glu
50 55 60
Arg Arg Glu Glu Glu His Ala Lys Leu Met Lys Ala Leu Lys Lys Cys
65 70 75 80
Lys Glu Glu Lys Gln Glu Ala Gln Lys Leu Met Asn Glu Val Gln Glu
85 90 95
Arg Leu Glu Glu Glu Glu Lys Leu Cys Gln Ala Ser Ser Ile Gly Ser
100 105 110
Trp Asp Gly Cys Arg Pro Cys Leu Glu Ser Asn Cys Ile Arg Phe Tyr
115 120 125
Thr Ala Cys Gln Pro Gly Trp Ser Ser Val Lys Ser Met Met Lys Gln
130 135 140
Phe Leu Lys Lys Ile Tyr Arg Phe Leu Ser Ser Gln Ser Glu Asp Val
145 150 155 160
Lys Asp Pro Pro Ala Ile Glu Gln Leu Thr Lys Glu Asp Leu Gln Val
165 170 175
Val His Ile Glu Asn Leu Phe Ser Gln Leu Ala Val Asp Ala Lys Ser
180 185 190
Leu Phe Asn Met Ser Phe Tyr Ile Phe Lys Gln Met Gln Gln Glu Phe
195 200 205
Asp Gln Ala Phe Gln Leu Tyr Phe Met Ser Asp Val Asp Leu Met Glu
210 215 220
Pro Tyr Pro Pro Ala Leu Ser Lys Glu Ile Thr Lys Lys Glu Glu Leu
225 230 235 240
Gly Gln Arg Trp Gly Ile Pro Asn Val Phe Gln Leu Phe His Asn Phe
245 250 255
Ser Leu Ser Val Tyr Gly Arg Val Gln Gln Ile Ile Met Lys Thr Leu
260 265 270
Asn Ala Ile Glu Asp Ser Trp Glu Pro His Lys Glu Leu Asp Gln Arg
275 280 285
Gly Met Thr Ser Glu Met Leu Pro Glu Gln Asn Gly Glu Met Cys Glu
290 295 300
Glu Phe Val Lys Asn Leu Ser Gly Cys Leu Lys Phe Arg Lys Arg Cys
305 310 315 320
Gln Lys Cys His Asn Tyr Leu Ser Glu Gly Ser Ser Lys Asp Ser Trp
325 330 335
Arg Arg Phe Phe Gln Ala Gly Glx
340

FIG.42B

```
Human          MKIKAEKNEG PSRSWWQLHW GDIANNSGNM KPPLLVFIVC LLWLKDSHCA
bovine         .......... .......... .........M KPPLLVFIVY LLRLRDCQCA
guinea         .......... .......... .........M KLPLLMFPVC LLWLKDCHCA
Rat HKNGTvar   .......... .......... .........M QPPLFVISVY LLWLKYCDSA
Rat HKNGD9     .......... .......... .........M QPPLFVISVY LLWLKYCDSA
Rat HKNGCvar   .......... .......... .........M QPPLFVISVY LLWLKYCDSA

Human          PTWKDKTAIS ENLKSFSEVG EIDADEEVKK ALTGIKQMKI MMERKEKEHT
bovine         PTGKDRTSIR EDPKGFSKAG EIDVDEEVKK ALIGMKQMKI LMERREEEHS
guinea         PTWKDKTAIS ENANSFSEAG EIDVDGEVKI ALIGIKQMKI MMERREEEHS
Rat HKNGTvar   PTWKETDATD GNLKSLPEVG EADVEGEVKK ALIGIKQMKI MMERREEEHA
Rat HKNGD9     PTWKETDATD GNLKSLPEVG EADVEGEVKK ALIGIKQMKI MMERREEEHA
Rat HKNGCvar   PTWKETDATD GNLKSLPEVG EADVEGEVKK ALIGIKQMKI MMERREEEHA

Human          NLMSTLKKCR EEKQEALKLL NEVQEHLEEE ERLCRESLAD SWGECRSCLE
bovine         KLMRTLKKCR EEKQEALKLM NEVQEHLEEE ERLCQVSLMG SWDECKSCLE
guinea         KLMKTLKKCK EEKQEALKLM NEVHEHLEEE ESLCQVSLAD SWDECRACLE
Rat HKNGTvar   KLMKALKKCK EEKQEAQKLM NEVQERLEEE EKLCQASSIG SWDGCRPCLE
Rat HKNGD9     KLMKALKKCK EEKQEAQKLM NEVQERLEEE EKLCQASSIG SWDGCRPCLE
Rat HKNGCvar   KLMKALKKCK EEKQEAQKLM NEVQERLEEE EKLCQASSIG SWDGCRPCLE

Human          NNCMRIYTTC QPSWSSVKNK IERFFRKIYQ FLFPFHEDNE KDLPISEKLI
bovine         SDCMRFYTTC QSSWSSMKST IERVFRKIYQ FLFPFHEDDE KELPIGEKFT
guinea         SNCMRFDTTC QPAWSSVKNK VEQFFRKIYQ FLFPLQE.ND RSGPVSKGVT
Rat HKNGTvar   SNCIRFYTAC QPGWSSVKSK MKQFLKKIYR FLSSQSE.DV KDPPAIEQLT
Rat HKNGD9     SNCIRFYTAC QPGWSSVKSK MKQFLKKIYR FLSSQSE.DV KDPPAIEQLT
Rat HKNGCvar   SNCIRFYTAC QPGWSSVKSK MKQFLKKIYR FLSSQSE.DV KDPPAIEQLT

Human          EEDAQLTQME DVFSQLTVDV NSLFNRSFNV FRQMQQEFDQ TFQSHFISDT
bovine         EEDVQLMQIE NVFSQLTVDV GFLYNMSFHV FKQMQQEFDL AFQSYFMSDT
guinea         EEDAQVSHIE HVFSQLSADV TSLFNRSLYV FKQLRREFDQ AFQSYFTSGT
Rat HKNGTvar   KEDLQVVHIE NLFSQLAVDA KSLFNMSFYI FKQMQQEFDQ AFQLYFMSDV
Rat HKNGD9     KEDLQVVHIE NLFDQLAVDA KSLFNMSFYI FKQMQQEFDQ AFQLYFMSDV
Rat HKNGCvar   KEDLQVVHIE NLFSQLAVDA KSLFNMSFYI FKQMQQEFDQ AFQLYFMSDV
```

FIG.43A

```
Human          DLTEPYFFPA FSKEPMTKAD LEQCWDIPNF FQLFCNFSVS IYESVSETIT
bovine         DSMEPYFFPA FSKEPAKKAH PMQSWDIPSF FQLFCNFSLS VYQSVSATVT
guinea         DVTEPFFFPS LSKEPAYRAD AEPSWAIPNF FQLLCNLSFS VYQSVSEKLI
Rat HKNGTvar   DLMEPYP.PA LSKEIIKKEE LGQRWGIPNF FQLFHNFSLS VYGRVQQIIM
Rat HKNGD9     DLMEPYP.PA LSKEITKKEE LGQRWGIPNF FQLFHNFSLS VYGRVQQIIM
Rat HKNGCvar   DLMEPYP.PA LSKEITKKEE LGQRWGIPNF FQLFHNFSLS VYGRVQQIIM

Human          KMLKAIEDLP KQDKAPDHGG LISKMLPGQD RGLCGELDQN LSRCFKFHEK
bovine         EMLKAIEDLS KQDKDSAHGG PSSTTWPVRG RGLCGEPGQN SSECLQFHAR
guinea         TTLRATEDPP KQDKDSNQGG PISKILPEQD RGSDGKLGQN LSDCVNFRKR
Rat HKNGTvar   KTLNAIEDSW EPHKELDQRG MTSEMLPEQN GEMCEEFVKN LSGCLKFRKR
Rat HKNGD9     KTLNAIEDSW EPHKELDQRG MTSEMLPEQN GEMCEEFVKN LSGCLKFRKR
Rat HKNGCvar   KTLNAIEDSW EPHKELDQRG MTSEMLPEQN GEMCEEFVKN LSGCLKFRKR

Human          CQKCQAHLSE DCPDVPALHT ELDEAIRLVN VSNQQYGQIL QMTRKHLEDT
bovine         CQKCQDYLWA DCPAVPELYT KADEALELVN ISNQQYAQVL QMTQHHLEDT
guinea         CQKCQDYLSD DCPNVPELYR ELNEALRLVS RSNQQYDQVV QMTQYHLEDT
Rat HKNGTvar   CQKCHNYLSE ECPDVPELHI EFLEALKLVN VSNQQYDQIV QMTQYHLEDT
Rat HKNGD9     CQKCHNYLSE GS..SKDSWR RFFQAG.... .......... ..........
Rat HKNGCvar   CQKCHNYLSE ECPDVPELHI EFLEALKLVN VSNQQYDQIV QMTQYHLEDT

Human          AYLVEKMRGQ FGWVSELANQ APETEIIFNS IQVVPRI..H EGNISKQDET
bovine         TYLMEKMREQ FGWVTELASQ TPGSENIFSF IKVVPGV..H EGNFSKQDE.
guinea         TLLMEKMREQ FGWVSELAYQ SPGAEDIFNP VKVMVALSAH EGNSSDQDD.
Rat HKNGTvar   IYLMEKMQEQ FGWVSQLASH NPVTEDIFNS TKAVPKI..H GGDSSKQDEI
Rat HKNGD9     .......... .......... .......... .......... ..........
Rat HKNGCvar   IYLMEKMQEQ FGWVSQLASH NPVTEDIFNS TKAVPKI..H GGDSSKQDEI

Human          MMTDLSILPS SNFTLKIPLE ESAESSNFIG YVVAKALQHF KEHFKTW
bovine         KMIDISILPS SNFTLTIPLE ESAESSDFIS YMLAKAVQHF KEHFKSW
guinea         TVVPSSLLPS SNFTLSSPLE KSAGNANFID HVVEKVLQHF KEHFKTW
Rat HKNGTvar   MVDSSSILPS SNFTVQNPPE EGAESSNVIY YMAAKVLQHL KGCFETW
Rat HKNGD9     .......... .......... .......... .......... .......
Rat HKNGCvar   MVDSSSILPS SNFTVQNPPE EGAESSNVIY YMAAKVLQHL KGCFETW
```

FIG.43B

cctgtagtcc cagctacgcg agaggctgag gcagcagaat tacttgaacc caggaggcgg 60
aggttgcagt gagccgagat cgcgccactg cactccagcc tgggtgagag agcgagactc 120
tgtctcaaaa aaaaaaaaaa aagaccgcca gggctcaaac aaaaaacctc ggaaaagccc 180
tggcggtctt ttttttttt ttttttttt tttttgggga cagtcttgct ctgtcgccca 240
ggctggagta caatggtcgg atcttggctc actgcaacct ctgcctccca ggttcaagca 300
attcttctgc ctcagcctcc caagtagcca ccacgcccag ctaatttttg tacttttagt 360
agagacgggg gtttcaccat gttgtccagg ctggtcttga actcctgacc tcaggtgatc 420
cacccgcctc ggcccccaa agtactagga ttacaggcgt gagccaccgc gtccagcgcc 480
ctggcggttt ttaatcaagt agaaaagctg cattatacca cttgcttcgg ttgcttcagt 540
gagaacgaag aaatggaaat gcaaatccct tattagttgt aggaaacaga tctcaaacag 600
cagttttgtt gacaagaccg caggaaaacg tgggaactgt gctgctggct tagagaaggc 660
gcggtcgacc agacggttcc caaagggcgc agtccttccc agccaccgca cctgcatcca 720
ggttcccggg tttcctaaga ctctcagctg tggccctggg ctccgttctg tgccacaccc 780
gtggctcctg cgtttccccc tggcgcacgc tctctagagc gggggccgcc gcgaccccgc 840
cgagcaggaa gaggcggagc gcgggacggc cgcgggaaaa ggcgcgcgga aggggtcctg 900
ccaccgcgcc acttggcctg cctccgtccc gccgcgccac ttggcctgcc tccgtcccgc 960
cgcgccactt cgcctgcctc cgtcccccgc ccgccgcgcc atgcctgtgg ccggctcgga 1020
gctgccgcgc cggcccttgc cccccgccgc acaggagcgg gacgccgagc cgcgtccgcc 1080
gcacggggag ctgcagtacc tggggcagat ccaacacatc ctccgctgcg gcgtcaggaa 1140
ggacgaccgc acgggcaccg gcaccctgtc ggtattcggc atgcaggcgc gctacagcct 1200
gagaggtgac gccgcgggcc cctgcgggac gggtggcggg aaggagggag gcgcggctgg 1260
ggagagcgct cgggagctgc cgggcgctgc ggaccccgtt tagtcctaac ctcaatcctg 1320
ccagggaggg gacgcatcgt cctcctcgcc ttacagacgc cgaaacggag ggtcccatta 1380
gggacgtgac tggcgcgggc aacacacaca gcagcgacag ccgggaggta agccgcgtcc 1440
cagcggctcc gcggccgggc tcgcagtcgc cccagtgatg ccgtggcccc cgaggcgggc 1500
gtcatcgggc agcgtttgcc cagtgctgga gggttaggga gagctgcctg ggcttgaccg 1560
cgcgccggtc tcaaagtcct ggctttggcc cctcctccgt tttcccctgt ggaccattcc 1620
gcttcgcagc gttttcaaaa actggagcga aagtgatgtg ggcggggcaa aggcggcggg 1680
aagaggacag cactgaagct ggcgcgggaa cttggtttcc tggtggcctc ccatccaatc 1740
cccacgaacc agctttcctc ttaaaccttg aaaagagaaa ttcgggagtt cgagttctta 1800
gtcgtccttt cctctttcct ttccgacagg agcaccccag gcaaaaaatg tctcgcgggt 1860
cattggcgcc aggctttcag gggacagtgg ggcggggcgg ggtgggcaca ggacgttagg 1920
cagccgttgg ccctccctaa ggccacaccg tcctgccgtc ctggatcctg cgccagctgc 1980
gcgggggagg ggactcgaag gtgtgtgagc caggggctga ccttgaccgc tcagataaat 2040
ggagcgcagc cttgacacag gggtggaggt ggttttgaat ggggaaaccc attcgtggtg 2100
aagcagattc actgtagcta gcggaaaagc cctccggccc acggacccat ctagagacga 2160
atacatagca gctgctgtgg ctgattggcg tgggacagcg tggggagttt tgtctgagga 2220
gagggatcca cttttctgca gctccaagcc caggggcctt tgatgagcca tagacctcat 2280
ttttaaccca cctttctgct tagacattga gcaagttact tctcatatag cttccctata 2340
tgttaaaaat ggagaaaata atgcttagta ggcaattctg ataaaagcag gtgcttgcaa 2400
aaatctctct gttgtctgaa tataaactgt accacaagcg agtgcggatg aacgaggact 2460
gcatttaaag ataagttttt acactttcat ttctctgtgg ctcgacactt ctgatgcctc 2520
ccttttgtt cctgggacac atgcttggtg ttgtcttcac acctttgtga caggattagc 2580
actagtgggc agtggatgat agctcctcct ccctttgcc acatgttcat ccctgccctc 2640
gccaccatct cactgtgtgg aattcctgtg tccactggtc accggggcac agaagtgctg 2700
tctcagcctg aatcgggcca ctgatgggac ttgcagcctg ggagctccac cgtgatctct 2760
ggcccacttt gcgggagtct aggctttctg gatgctccag gcctcacgtc ccagggcagt 2820
tttcttccct gaagaaagtt ggatggcatg atctgtcttc ccatcttgaa accgtatggc 2880
aaattgtttt tcagatgaat tccctctgct gacaaccaaa cgtgtgttct ggaagggtgt 2940
tttggaggag ttgctgtggt ttatcaaggt aaagaagtcg ctgctattag aagtcagtag 3000

FIG.44A

tctgttctca acacagcagc cagtgagatc ctttcaaaac tcaaagcagc caggtgtggt 3060
ggctcacgcc tgtaatccca ccgctttggg aggctgagtc agatcacctg aggttaggaa 3120
tttgggacca gcctggccaa catggcgaca ccccagtctc tactaataac acaaaaaatt 3180
agccaggtgt gctggtgcat gtctgtaatc ccagctactc aggaggctga ggcatgagaa 3240
ttgctcacga ggcggaggtt gtagtgagct gagatcgtgg cactgtactc cagcctggcg 3300
acagagggag aacccatgtc aaaaacaaaa aaagacacca ccaaaggtca aagcatatca 3360
ttcctcaccc tcaagccctt agtggctcca tttcactcag taagagccac ggtccttatg 3420
gtgtccgttt ttcagctctg accttagctg ctgctctctg caccaccctg ctgttcttgt 3480
gagtttttga gcacaccggg acatccccac tccctggaac cttcttcccc cacacttggc 3540
ttcttccttt gagtctctac tccactcggg caagccttcc tagacctcct gatttaaaac 3600
tgtgactctc ccccaacctc cttggtgttt ctccgtagac gaacatcacc atctgatgta 3660
tgtcagcctt tcccttcccc tgttagaagg gggacagcag gtagtaaaag tgaaatgtgc 3720
tgtaagcttt atgagggcag aggatttgtt tctcgtgttc actgttgtat cgccagggcc 3780
tcaaacacag cctgccacat agtaggagtc aacatatatt gatcactaaa tgtagatacc 3840
acctgtgttc ccatgttcat ataaattcta gaagagtctc ttcagtaaca aggtgaaccc 3900
cttccagagg gctgagtagg tacctcaggc cggggccaga gtgctgtgaa gacagcagca 3960
gcccagacca agcttctctg tgttccgtgt cctggtctag aaccagcgat gttctttctg 4020
accagtgctt tttggaaggt ggctgaggtc tgggctcagg tctgggccat actagaagct 4080
gggatccctt ctatagagca cttggtatgg cttgtatggt cttggggcaa gccagaccca 4140
agccctctta tcccatttta gaaagggctt caatttggat ccagccccag gtctgcctta 4200
gctctgtatt cttggggtat tttgttctgt attggcctat cttgactaac aatgagcctt 4260
ggatttgaaa catatcatca gaaacctcag aagacaacat tcttaaactg gctagagcct 4320
ggtctgaatg gatgaaaagg agagactttt gaagcaatat gtaaaagatt gagaaatgat 4380
ttgttggaaa tttctcaatt ggagaaattt ctttgatttg ttggaaattt ctttgattct 4440
ttctcaatca aagaaaatcg ggacaaactc aacaatagaa agggaggaag caagatactc 4500
agaaataaaa tgcattcccc tgtttcaact taatgcttca attcaggatt ctaaggaatc 4560
cttgccagga atgtcagact caccttgata gttggagtta ctccattggt gactcgatca 4620
aatacaggag ttgaggcacc tgcactgtaa aatactgatt agtctgatca ttaggaatat 4680
cctgtatgcc aggtagaaga tacattgaac agattgcatg taggcattaa attcattttg 4740
gggtattaca tatagacaac acatttcatt aagaaacata aaactgtcag atcggtggaa 4800
tacttaaaag cacttggagg tgtttagcct aaaaagctta gttgagggga atggaagaaa 4860
agatctggga gggtggttcc aaagaaggga tcagactatc ctaaagccct caggaatctg 4920
ggctgggacc acctacttaa agataggatg ggcagctggg tgtggtggct cacgcctgta 4980
atcccagcac ttcgggaggc cgaagcgggc ggatcacctg aggtcaggag ttcgaggcca 5040
gcctgaccaa catggagaaa cgctgtctct actaaaaata caaaattagc tgggtgtagt 5100
ggcgcatgcc tgtaatccca gctactcggg aggctgaggc aggggaatcg cttgaacctg 5160
ggaggtggag ggtgccgtga gccacgatcg cgccattgca ctccagcctg ggcaacaaga 5220
gcgaaactct caaaaaacaa aaaaaaggat gggttccata tgggtggtgt caagtgccca 5280
cctcctagca agtcagcagg ggccagaggc ccttgtaagt ggtgtctcgg ggggatcaac 5340
tgagatggct taagatttac ctggatgcct gctctgctct ccccatctct tccagggatc 5400
cacaaatgct aaagagctgt cttccaaggg agtgaaaatc tgggatgcca atggatcccg 5460
agacttttg gacagcctgg gattctccac cagagaagaa ggggacttgg gcccagttta 5520
tggcttccag tggaggcatt ttggggcaga atacagagat atggaatcag gtgaggagat 5580
agaacaatgc cttccatttc cgggtgccct tcctagcacg tgtttgctcc gttgttttag 5640
ataaggtctg ggggatgagt caatgtcaca ggagctgatg tatagctttg accttgtgag 5700
gggtggtgcc aggttgaagc cacaattaac gcctactgaa ggccgtttca catcttttt 5760
ttttttttt ttttaattat tatactttaa gttttagggt acatgtgcac aatgtgcagg 5820
ttagttacat atgtatacat gtgccatgct ggtgcgctgc accactaact caccatctag 5880
catcaggtat atctcccaat gctatccctc ccccctcctc ccaccccaca acatccccag 5940
agtgtgatgt tcccccttcct gtgtccatat gttctcgttg ttcgattccc actatgagtg 6000

FIG.44B

agaatatgcg gtgtttggtt ttttgttctt gcgatagttt actgagaatg atgatttcca 6060
tttcaccacg tccctacaga ggacatgaac tcatcatttt ttatggctgc atagtattcc 6120
atggtgtata tgtgccacat tttcttaatc cagtctatca tgttggacat ttgggttggt 6180
tccaagtctt tgcctattgt gaatagtgcc acaataaaca tacgtgtgca tgtgtcttta 6240
tagcagcatg atttaatagt cctttgggta tatacccagt aatgggatgg ctgggtcaaa 6300
tggtatttct agttctagat ccccgaggaa tcgccacact gacttccaca atggttgaac 6360
tagtttacag tcccaccaac agtgtcaaag tgtcctattt ctccacatcc tctccagcac 6420
ctgttgtttc ctgacttttt aatgattgcc attctaactg gtgtgagatg gtatctcatt 6480
gtggttttga tttgcgtttc tctgatggcc agtgatggtg agcatttttt catgtgtttt 6540
ttggctgcat aaatgtcttc ttttgagaag tgtctgttca tgtccttcgc ccacttttg 6600
atggggttgt ttttttctta taaatttgtt tgagttcatt gtagattctg gatattagcc 6660
ctttgtcaga tgagtaggtt gcaaaaatgt tctcccattt tgtgggttgc ctgttcactc 6720
tgatggtagt ttcttttgct gtgcagaagc tctttagttt aattagatcc catttgtcaa 6780
ttttggcttt tgttgccatt gcttttggca taggcatgaa gtccttgccc atgcctatgt 6840
cctgaatggt aatgcctagg ttttcttcta gggtttttat ggttttaggt ctaacgttta 6900
agtctttaat ccatcttgaa ttgatttttg tataaggtgt aaggaaggga tccagtttca 6960
gctttttaca tatggctagc cagttttccc agcaccattt attacatagg gaatcctttc 7020
cccattgctt gtttttctca ggtttgtcaa agatcagata gttgtagata tgcggcgtta 7080
tttctgaggg ctctgttctg ttccattgat ctatgtgtct gttttggtac cagtaccata 7140
ctgttttggt tactgtagcc ttgtagtata gtttgaagtc aggtagcgtg atgcctccag 7200
ctttgttctt ttggcttagg attgacttgg cgatgcgggc tctttttgg ttccatatga 7260
actttaaagt agtttttcc aattctgtga agaaagtcat tggtagcttg atggggatgg 7320
cattgaatct ataaattacc ttgggcagta tggccatttt cacgatattg attcttccta 7380
cccatgagca tggaatggtc ttccatttct ttgtatcctc ttttatttca ttgagcagtg 7440
gtttgtagtt ctccttgaag aggtccttca catccctttt aaggtggatt cctaggtatt 7500
ttattctctt tgaagcaatt gtgagtggaa gttcactcat gatttggctc tctgtttgtc 7560
tgttattggt gtataagaat gcttgtgatt tttgcagatt gattttatat cctgagactt 7620
tgctgaagct gcttatcagc ttaaggagat tttgggctga gacaatgggg ttttctagat 7680
atacaatcat gtcgtctgca aacagggaca atttgacttc ctcttttcct aattgaatac 7740
cctttatttc cttctcctgc ctaattgccc tggccagaac ttccaacact atgttgaata 7800
ggagtggtga gagagggcat ccctgtcttg tgccagtttt caaagggaat gcttccagtt 7860
tttgcccatt cactatgata ttggctgtgg ctttgtcata gatagctctt attattttga 7920
aatatgttcc atcaatacct aatttattga gagtttttag catgatgtgt tgttgaattt 7980
tgtcaaaggc tttttctgca tctattgaga taatcatgtg gtttttgtct ttggatctgt 8040
ttatatgctg gattacattt attgatttgc gtatattgaa ccagccttgc atcctaggga 8100
tgaagcccac atgatcatgg tggataagct ttttgatgtg ctgctggatt cggtttgcca 8160
gtattttatt gaggattttt gcatcaatgt tcatcaagga tattggtcta aaattctctt 8220
ttttggtgtg tctctgccca gctttggtat caggatgatg ttggcttcat aaaatgagtt 8280
agggaggatt ccctcttttt ctattgattg gaatagtttc agaaggaatg gtaccagttc 8340
ctctttgtac ctctggagaa ttcggctgtg aatccatctg gtcctggact ctctttggtt 8400
ggtaagctat tgattattgc cacaatttca gctcctgtta ttggtctatt cagagattca 8460
acttcttcct ggtttagtct tgggagagtg tatgtgtcaa ggaatttatc catttcttct 8520
agattttcta gtttatttgc gtagaggtgt ttgtagtaat ctctgatggt agtttgtatt 8580
tctgtgggat cggtggtgat atccccttta tcatttttta ttgcgtctat ttgattcttc 8640
tctttttctt tattagtctt gctagcggtc tataaatttt gttgatcctt tcaaaaaacc 8700
agctcctgga ttcattaatt ttttgaaggg ttttttgtgt ctctatttcc ttcagttctg 8760
ctctgatttt agttatttct tgccttctgc tagcttttga atatgtttgc tcttgctttt 8820
ctagttcttt taattgtgat gttagggtgt caattttgga tctttcctgc tttctcttgt 8880
gggcatttag tgctataaat ttccctctac acactgcttt gaatgtgtcc cagaggttct 8940
ggtatgttgt gtctttgttc ttgttggttt caaagaacat ctttatttct gccttcattt 9000

FIG.44C

cgttatgtac ccagtagtca ttcaggagca ggttgttcag tttccatgta gttgagcagt 9060
tttgagtgag attcttaatc ctgagttcta gtttgattgc actgtggtct gagagatagt 9120
ttgttataat ttctgttctt ttacatttgc tgaggagagc tttacttcca actatgtggt 9180
cggttttgga ataggtgtgg tgtggtgctg aaaaaaatgt atattctgtt gatttgggat 9240
ggagttctgt agatgtctat taggtctgct tggtgcagag ctgagttcaa ttcctgggta 9300
tccttgttga ctttctgtct cgttgatctg tgtactgttg acagtgggtg ttaaagtctc 9360
ccattattaa tgtgtggagt ctaagtctct ttgtaggtca ctcagatgat tggcacttac 9420
tgggcgcttg gcactttcca tactgtgtca tcggcagata gctgcatggt tggtgttcgt 9480
gctggggaat gggaagttca tcggtgggac aaggacaaaa tgcccccatt gctttgttgt 9540
ggctttaatc tccctttcga ggctgagcca cagcgtgctg taggtggcgc tgctgtgaag 9600
cgcagtacca gggtcacact ccactcccag ctctgcagag gtggagaaag aatgaaacat 9660
ctcactcctg gacttccact ttcctgtcac tgttggtgtc acctcttact ggatgtcaca 9720
gagcccagcc cctcccacct gtgcctagga aaagcagatg ccaccttgga atgtggggtt 9780
tgtgtgtgca atttactagc tgggcagaga ccagcaacct ggagagcagg tgtctcgtct 9840
aaggggacag tcacatttca cctccagcca cctggaggaa tttgggcctg gtgatgtcag 9900
aattcttcaa taaaagccta aaatctatat tttatgtgcg gtcatgagat ctgttaaatg 9960
ttagcaactt caggaagttt aaaaatgctg tgtggaccta gaataggcaa gttcttaaag 10020
gcagaaagtg gaatgctagt ttccagggac tggggaacag ggaggaatgg ggagttcatg 10080
tttaatgggc acagaggttt tgttagggat gacgaaaaag ttcgggagat ggtgatggtg 10140
atggagatgg tgatggtgat ggagatggtg atggtgatgg tgatggtgat gggtgatggt 10200
gatggtgatg gtgatggtga tggagatggt gatggtgatg gtgatggaga tggtgatggt 10260
gatggtgatg gtgatggaga tggtgatggt gatggagatg gtgatggtga tggtgatgga 10320
gatggtgatg gtgatggtga tggtgatggt gatggtgatg gtgatggaga tggagatggt 10380
gatggtgatg gttgcctaac atcaggaacg tgcttaatgc ttctgaattg cacacaaaaa 10440
tggcaagttt aatattatgt gtactttatc acaatgaaaa aagctgctgc gtgggccaag 10500
ttacttgtgc aggtaatgtt ctgcaggtgg ttgcctgcac ctcagttgta gggtgtccgt 10560
aggatgtgag gccagtcccc gggcttaatg atgctttaaa tcctgcctag tattcaatta 10620
tttcttgtcg cttaaaaggc ctaataaaat tatggtctta gtttacagtg gtatgaatgc 10680
ttagctgttg gattttagta ggaaagttcg tccctttttg tttttaattt tgttttacag 10740
attcacagga attttttttt ttttttttt ttttttttt taatgcacag aaagtttccc 10800
tggactctct acccagtttc cccagtgata atatcttggg taacatcctg tatacattca 10860
cattggtgca ttcctcagag ttgtcagatt ttgctagttt tacgtgcact tgtgtatgtg 10920
tgtatttgca attttagcac gtgtagactc ttgtaaccac tacaatcaag ttacagaact 10980
acactaccaa ggttcatctt tttaaaatct ttgatgttac ctttttttgga acagtgacca 11040
tgagaggact ttcctcccaa aattttgaaa actactgaac cagaatatag tctgacacta 11100
ataggtagaa atttaaccaa aggagattat gaagctctgc acttgagtta acaaaatcac 11160
ttctcagctt ccagttccat ctcagaagga aggaaaaggg attaaaaatc cagagaccag 11220
aaaatgggag caaagtacaa ggtggtgtaa tcattacaga ggtttcctga tgtttccaag 11280
tcagtcgtgt gttgagctgc taaactctaa agtaatttta ggtggaatgt tggaaacatg 11340
ctgctgaggt gatagaaagg aatccatggt cctctgttag ttggaaagta tatggaatac 11400
tatattctac ataagataca atactctctg tgagacaagg ataaagtaga ttttgtcagt 11460
gaaattgtga caagaatcgc tgatgggttt agagcctaag tttgcgagga gcactggaag 11520
aaattaagat tgttgagatt ggaaagggtt agctatgggg gaacaggagg aggtgactcc 11580
atgacagacc aaatattcaa aggactgtgt agaagaggaa aaagactttg ttagggctcc 11640
agaggacaga gccaggagtc agacagggcc ttgaactcaa cccaccgaga tctgcaaact 11700
ttgcaggatg caccagatgt cttgtagcca tgggtcaagg ggggaccctg ggtaagagac 11760
tgtaatagat gacctctaag gccatctcat gacatgtgtg attaatgtat gtacctgtcc 11820
tctcttttg acaattctac agattattca ggacagggag ttgaccaact gcaaagagtg 11880
attgacacca tcaaaaccaa ccctgacgac agaagaatca tcatgtgcgc ttggaatcca 11940
agaggttgaa agaaccccgt cgtcttcatt tatactaacc atactcttag agggaagcaa 12000

FIG.44D

tctggttttg tgcagaggca ctgagggagg caggaccctg ggcaacttcc cccagccaca 12060
tggttgtgtg acgttgggca agtcacattt tgctgcactt tcaccttcag atcatgaggt 12120
tgggcccaga ggatttttt ttttttttt ttttttgaga cagagttttg ctctgttgcc 12180
caggctggaa tgcaacggcg tgatcttggc tcactgtaac ctctgcctcc tgggttcgag 12240
tgattctcct gcctcagcct ccaagtagct gggattacag catgtgccac catgcctggc 12300
taattttgta ttttagtag agacgggttc acatgttggt caggctggtc ttgactcctg 12360
accctcagat gatctgcctt gcctcagcct cccaaccgag tgatcttaag ttgtgtatta 12420
tactcattct tacacaaaaa gggctttaaa tgcctagaaa ctacatgaag atgttaacat 12480
tttaaatgga agcagatgaa gttccagctc gctgccacct cactaacatt tttaacaatt 12540
atattgtaaa attcaactct accagggtgt agagccaggt gtggtggctc acacctgtaa 12600
ttccaacaac tccagaggcc aaggcgagag gatcatttga acccacggaa tttgaggctg 12660
tagtgagtca tgatcacgcc attgcactcc atcctgggca acagagtgag accctgaata 12720
tttaaaaaca acaacaacaa caaaactcta tcaggatatc ataagtactt agagtgaaat 12780
acttgcatct gtaatagaga cttatttttt ttttttttga gacacagtct caccctgttg 12840
cccaggctgg agtgcagtgg tttgatctcc gctcacggca acctccatct cccaggttca 12900
agtgagttcc cattcctcag ccccagagct gggaccacag gcgcgcgaat ttttgtattt 12960
ttagcagaga cggggtttca ctatgttggc caggctagtc tcaaactcaa gttggcctca 13020
agtgatctgc ccaccctggc gtcccagtgt tgggatttca ggcatgagcc actgtgcctg 13080
gccatgtaat agagactttt aatataggag ggtgtaccag aagcaccagt ttcctgtggc 13140
aaacagaatt attcctgctg tatttgtaat ttggtgccac gaggtagccc agatcccttc 13200
agctctgatg gaagagcatt gcttcagccg taaatggaca cctgcagaaa ccttgcaccg 13260
atggatagtc tccctcagct ccgtgccatc gctgcagggg ctgttatgga catcactgca 13320
gcccagtggc tctctctcct ggtctccacc atatgagttg gcttctgttt ctctcctgtt 13380
ttactttgcc tttagctgtg gtctttcaaa ccaccatccc tccttatctt cctctgctgg 13440
ttcctcagat cttcctctga tggcgctgcc tccatgccat gccctctgcc agttctatgt 13500
ggtgaacagt gagctgtcct gccagctgta ccagagatcg ggagacatgg gcctcggtgt 13560
gcctttcaac atcgccagct acgccctgct cacgtacatg attgcgcaca tcacgggcct 13620
gaaggtgggc tgtctcggga agggtgactt gccagcctac cacatgagct cttcagttct 13680
ttaatatggg aaaacaaatt gcagagttta gtctctgatt agcttttaaa tttgatatgt 13740
gtaagtaaga catgaaccag cttttacttt gaaaccttcc ttttctggaa ggttttctgg 13800
ccctgtggta tatgcactaa cagatctata caggttgttt gtgatacagc ttctatggat 13860
cttctcaaaa gctatgctga ggttgggtat ggtggctcat gcctgtaatc ccagcacttt 13920
ggaagactga gacaggagca attgcttgag gtctggagtt caataccagc ctgggcaaca 13980
taacaagatg ctgttgctac aaaaaaatgg aaaagctaca ctaaattatt tttttaaaaa 14040
aagccttgcg gtgtctgcat attctaatgt ttttaaatga tgttttaaag aattgaaact 14100
aacatactgt tctgctttct cccggtttat agccaggtga ctttatacac actttgggag 14160
atgcacatat ttacctgaat cacatcgagc cactgaaaat tcaggtaaga attagatgtt 14220
atactttgg gtttggtacc ttctcttgat aaaaggttga ctgtggaaca ggtatctgct 14280
caatgctgtg tccaagataa agatgactgc tccaaatgtg gggcttcagt ttagggagaa 14340
gtggtgggca ggtgggcagg acaaggcagg catctgcctc agcaaccatg gcacttaact 14400
tgtcaggtgc tgtgaggtac taagcaccag taccagagag ggaagagcca cattcaagcc 14460
aggggattgt ccaaaaggag gcattttaac tcattttaac ttgaaggaga attgaagtgc 14520
aaatgttttt ccttttcttt ttttttgaga tggagtcttt ctctgtcggc caggctggag 14580
tgtgccgtgg tgcgatctca gctcactgca acctccacct cccgggttca agcaattctt 14640
ctgcctcagc ctcccaggta gctgggatta caggcacatg ccaccacacc cagctaattt 14700
tttgtattat tagtagagat ggggtttcgt catgttggcc aggctgatct caaactcctg 14760
acttcaagtg taccacctgc ctcagcctcc gaaagttctg gaattacagg cataagccac 14820
caccctggcc ataaatattt tttgttaatt ttacattaag tacaatattt aggtccaaac 14880
ttcaaaagtc tgttgaaatc cctgaagtta tagcagccaa caattgatat gaaatggcaa 14940
taaaaatgta agttcatctg cttcatgagc cttaaggaaa aaaactcaga accagacact 15000

FIG.44E

ttttagcccc ttccaggtta gatccaggtt ttaaaagtta ttcctttgag ggagtttggc 15060
tgcttttgag tggaggtgac ttcaggctta ttctctctgg ctctctgctc tggtcatttt 15120
tagacatagt aataggttgt gacctgtctt cacatcctaa ttgccactgt ctgttcatcc 15180
caggaatcct ggctttcatc cctttctgtt cactgtccat gcatgtcatc tttccttctt 15240
tctgccaggg accagatggg ttagggattg tgaattcaag taaacgtaga gctactatga 15300
gttacagatt gactgtgttc ctgtctttaa taaatttgcc aagagtggtt ataagaactt 15360
acacctgatg aggcaccagg ctcctgatgc tgtgtaatgt cacaaaatac ccctcactct 15420
cgatctgtgc aagagaacag ctggttgcgc tccaatcatg ttacataacc tacgcgaagg 15480
tatcgacagg atcatactcc tgtaaaatag aactttgttg atcacatcct gtgtacttgt 15540
ttcacggaca tgaggagcaa ttacaacagg tcgtacaatt atggcaaaat aatggcctta 15600
ttttgttttt agcttcagcg agaacccaga cctttcccaa agctcaggat tcttcgaaaa 15660
gttgagaaaa ttgatgactt caaagctgaa gactttcaga ttgaagggta caatccgcat 15720
ccaactatta aaatggaaat ggctgtttag ggtgctttca aaggagctcg aaggatattg 15780
tcagtcttta ggggttgggc tggatgccga ggtaaaagtt ctttttgctc taaaagaaaa 15840
aggaactagg tcaaaaatct gtccgtgacc tatcagttat taatttttaa ggatgttgcc 15900
actggcaaat gtaactgtgc cagttctttc cataataaaa ggctttgagt taactcactg 15960
agggtatctg acaatgctga ggttatgaac aaagtgagga gaatgaaatg tatgtgctct 16020
tagcaaaaac atgtatgtgc atttcaatcc cacgtactta taaagaaggt tggtgaattt 16080
cacaagctat ttttggaata tttttagaat attttaagaa tttcacaagc tattccctca 16140
aatctgaggg agctgagtaa caccatcgat catgatgtag agtgtggtta tgaactttaa 16200
agttatagtt gttttatatg ttgctataat aaagaagtgt tctgcattcg tccacgcttt 16260
gttcattctg tactgccact tatctgctca gttccttcct aaaatagatt aaagaactct 16320
ccttaagtaa acatgtgctg tattctggtt tggatgctac ttaaaagagt atattttaga 16380
aataatagtg aatatatttt gccctatttt tctcatttta actgcatctt atcctcaaaa 16440
tataatgacc atttaggata gagtttttt ttttttttt taaactttta taaccttaaa 16500
gggttatttt aaaataatct atggactacc attttgccct cattagcttc agcatggtgt 16560
gacttctcta ataatatgct tagattaagc aaggaaaaga tgcaaaacca cttcggggtt 16620
aatcagtgaa atatttttcc cttcgttgca taccagatac ccccggtgtt gcacgactat 16680
ttttattctg ctaatttatg acaagtgtta aacagaacaa ggaattattc caacaagtta 16740
tgcaacatgt tgcttatttt caaattacag tttaatgtct aggtgccagc ccttgatata 16800
gctatttttg taagaacatc ctcctggact ttgggttagt taaatctaaa cttatttaag 16860
gattaagtag gataacgtgc attgatttgc taaaagaatc aagtaataat tacttagctg 16920
attcctgagg gtggtatgac ttctagctga actcatcttg atcggtagga ttttttaaat 16980
ccatttttgt aaaactattt ccaagaaatt ttaagccctt tcacttcaga aagaaaaaag 17040
ttgttggggc tgagcactta attttcttga gcaggaagga gtttcttcca aacttcacca 17100
tctggagact ggtgtttctt tacagattcc tccttcattt ctgttgagta gccgggatcc 17160
tatcaaagac caaaaaaatg agtcctgtta acaaccacct ggaacaaaaa cagattttat 17220
gcatttatgc tgctccaaga aatgctttta cgtctaagcc agaggcaatt aattaatttt 17280
ttttttttg acatggagtc actgtccgtt gcccaggctg cagtgcagtg gcgcaatctt 17340
ggctcactgc aacctccacc tcccaggttc aagtgattct cctgcctcag cctcccatgt 17400
agctgggatc acaggcacct gccaccatgc ccggctaatt ttttgtattt tttgtagaga 17460
cagggtttca ccatgttggc caggctggtc tcaaacacct gacctcaaat gatccacctg 17520
cctcagcctc ccaaagtgtt gggattacag gcgtaagcca ccatgcccag ccctgaatta 17580
atatttttaa aataagtttg gagactgttg gaaataatag ggcagaggaa catattttac 17640
tggctacttg ccagagttag ttaactcatc aaactctttg ataatagttt gacctctgtt 17700
ggtgaaaatg agccatgatc tcttgaacat gatcagaata aatgccccag ccacacaatt 17760
gtagtccaaa ctttttaggt cactaacttg ctagatggtg ccaggttttt ttgcacaagg 17820
agtgcaaatg ttaagatctc cactagtgag gaaaggctag tattacagaa gccttgtcag 17880
aggcaattga acctccaagc cctggccctc aggcctgagg attttgatac agacaaactg 17940
aagaaccgtt tgttagtgga tattgcaaac aaacaggagt caaagcttgg tgctccacag 18000

FIG.44F

```
tctagttcac gagacaggcg tggcagtggc tggcagcatc tcttctcaca ggggccctca   18060
ggcacagctt accttgggag gcatgtagga agcccgctgg atcatcacgg gatacttgaa   18120
atgctcatgc aggtggtcaa catactcaca caccctagga ggagggaatc agatcggggc   18180
aatgatgcct gaagtcagat tattcacgtg gtgctaactt aaagcagaag gagcgagtac   18240
cactcaattg acagtgttgg ccaaggctta gctgtgttac catgcgtttc taggcaagtc   18300
cctaaacctc tgtgcctcag gtccttttct tctaaaatat agcaatgtga ggtggggact   18360
ttgatgacat gaacacacga agtccctctg agaggttttg tggtgccctt taaaagggat   18420
caattcagac tctgtaaata tccagaatta tttgggttcc tctggtcaaa agtcagatga   18480
atagattaaa atcaccacat tttgtgatct atttttcaag aagcgtttgt atttttttcat   18540
atggctgcag cagctgccag gggcttgggg ttttttttggc aggtagggtt gggagg       18596
```

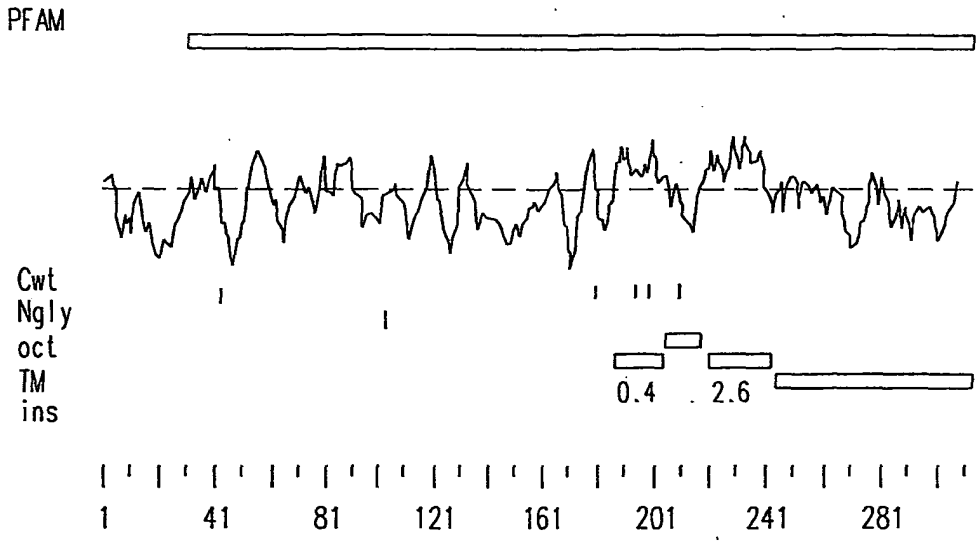
FIG.44G

```
gggggggggg ggaccacttg gcctgcctcc gtcccgccgc gccacttggc ctgcctccgt      60
cccgccgcgc cacttcgcct gcctccgtcc cccgcccgcc gcgccatgcc tgtggccggc     120
tcggagctgc cgcgccggcc cttgcccccc gccgcacagg agcgggacgc cgagccgcgt     180
ccgccgcacg gggagctgca gtacctgggg cagatccaac acatcctccg ctgcggcgtc     240
aggaaggacg accgcacggg caccggcacc ctgtcggtat tcggcatgca ggcgcgctac     300
agcctgagag atgaattccc tctgctgaca accaaacgtg tgttctggaa gggtgttttg     360
gaggagttgc tgtggtttat caagggatcc acaaatgcta aagagctgtc ttccaaggga     420
gtgaaaatct gggatgccaa tggatcccga gacttttgg acagcctggg attctccacc      480
agagaagaag gggacttggg cccagtttat ggcttccagt ggaggcattt tggggcagaa     540
tacagagata tggaatcaga ttattcagga cagggagttg accaactgca aagagtgatt     600
gacaccatca aaaccaaccc tgacgacaga agaatcatca tgtgcgcttg gaatccaaga     660
gatcttcctc tgatggcgct gcctccatgc catgccctct gccagttcta tgtggtgaac     720
agtgagctgt cctgccagct gtaccagaga tcgggagaca tgggcctcgg tgtgcctttc     780
aacatcgcca gctacgccct gctcacgtac atgattgcgc acatcacggg cctgaagcca     840
ggtgacttta tacacacttt gggagatgca catatttacc tgaatcacat cgagccactg     900
aaaattcagc ttcagcgaga acccagacct ttcccaaagc tcaggattct tcgaaaagtt     960
gagaaaattg atgacttcaa agctgaagac tttcagattg aagggtacaa tccgcatcca    1020
actattaaaa tggaaatggc tgtttagggt gctttcaaag gagcttgaag gatattgtca    1080
gtctttaggg gttgggctgg atgccgaggt aaaagttctt tttgctctaa aagaaaaagg    1140
aactaggtca aaaatctgtc cgtgacctat cagttattaa tttttaagga tgttgccact    1200
ggcaaatgta actgtgccag ttctttccat aataaaaggc tttgagttaa ctcactgagg    1260
gtatctgaca atgctgaggt tatgaacaaa gtgaggagaa tgaaatgtat gtgctcttag    1320
caaaaacatg tatgtgcatt tcaatcccac gtacttataa agaaggttgg tgaatttcac    1380
aagctatttt tggaatattt ttagaatatt ttaagaattt cacaagctat tccctcaaat    1440
ctgagggagc tgagtaacac catcgatcat gatgtagagt gtggttatga actttatagt    1500
tgttttatat gttgctataa taaagaagtg ttctgc                              1536
```

FIG.45A

```
Met Pro Val Ala Gly Ser Glu Leu Pro Arg Arg Pro Leu Pro Pro Ala
 1               5                  10                  15
Ala Gln Glu Arg Asp Ala Glu Pro Arg Pro Pro His Gly Glu Leu Gln
            20                  25                  30
Tyr Leu Gly Gln Ile Gln His Ile Leu Arg Cys Gly Val Arg Lys Asp
        35                  40                  45
Asp Arg Thr Gly Thr Gly Thr Leu Ser Val Phe Gly Met Gln Ala Arg
    50                  55                  60
Tyr Ser Leu Arg Asp Glu Phe Pro Leu Leu Thr Thr Lys Arg Val Phe
65                  70                  75                  80
Trp Lys Gly Val Leu Glu Glu Leu Leu Trp Phe Ile Lys Gly Ser Thr
                85                  90                  95
Asn Ala Lys Glu Leu Ser Ser Lys Gly Val Lys Ile Trp Asp Ala Asn
            100                 105                 110
Gly Ser Arg Asp Phe Leu Asp Ser Leu Gly Phe Ser Thr Arg Glu Glu
        115                 120                 125
Gly Asp Leu Gly Pro Val Tyr Gly Phe Gln Trp Arg His Phe Gly Ala
    130                 135                 140
Glu Tyr Arg Asp Met Glu Ser Asp Tyr Ser Gly Gln Gly Val Asp Gln
145                 150                 155                 160
Leu Gln Arg Val Ile Asp Thr Ile Lys Thr Asn Pro Asp Asp Arg Arg
                165                 170                 175
Ile Ile Met Cys Ala Trp Asn Pro Arg Asp Leu Pro Leu Met Ala Leu
            180                 185                 190
Pro Pro Cys His Ala Leu Cys Gln Phe Tyr Val Val Asn Ser Glu Leu
        195                 200                 205
Ser Cys Gln Leu Tyr Gln Arg Ser Gly Asp Met Gly Leu Gly Val Pro
    210                 215                 220
Phe Asn Ile Ala Ser Tyr Ala Leu Leu Thr Tyr Met Ile Ala His Ile
225                 230                 235                 240
Thr Gly Leu Lys Pro Gly Asp Phe Ile His Thr Leu Gly Asp Ala His
                245                 250                 255
Ile Tyr Leu Asn His Ile Glu Pro Leu Lys Ile Gln Leu Gln Arg Glu
            260                 265                 270
Pro Arg Pro Phe Pro Lys Leu Arg Ile Leu Arg Lys Val Glu Lys Ile
        275                 280                 285
Asp Asp Phe Lys Ala Glu Asp Phe Gln Ile Glu Gly Tyr Asn Pro His
    290                 295                 300
Pro Thr Ile Lys Met Glu Met Ala Val
305                 310
```

FIG.45B



MPVAGSELPRRPLPPAAQERDAEPRPPHGELQYLGQIQHILRCGVRKDDRTGTGTLSVFG
MQARYSLRDEFPLLTTKRVFWKGVLEELLWFIKGSTNAKELSSKGVKIWDANGSRDFLDS
LGFSTREEGDLGPVYGFQWRHFGAEYRDMESDYSGQGVDQLQRVIDTIKTNPDDRRIIMC
AWNPRDLPLMALPPCHALCQFYVVNSELSCQLYQRSGDMGLGVPFNIASYALLTYMIAHI
TGLKPGDFIHTLGDAHIYLNHIEPLKIQLQREPRPFPKLRILRKVEKIDDFKAEDFQIEG
YNPHPTIKMEMAV

FIG.46

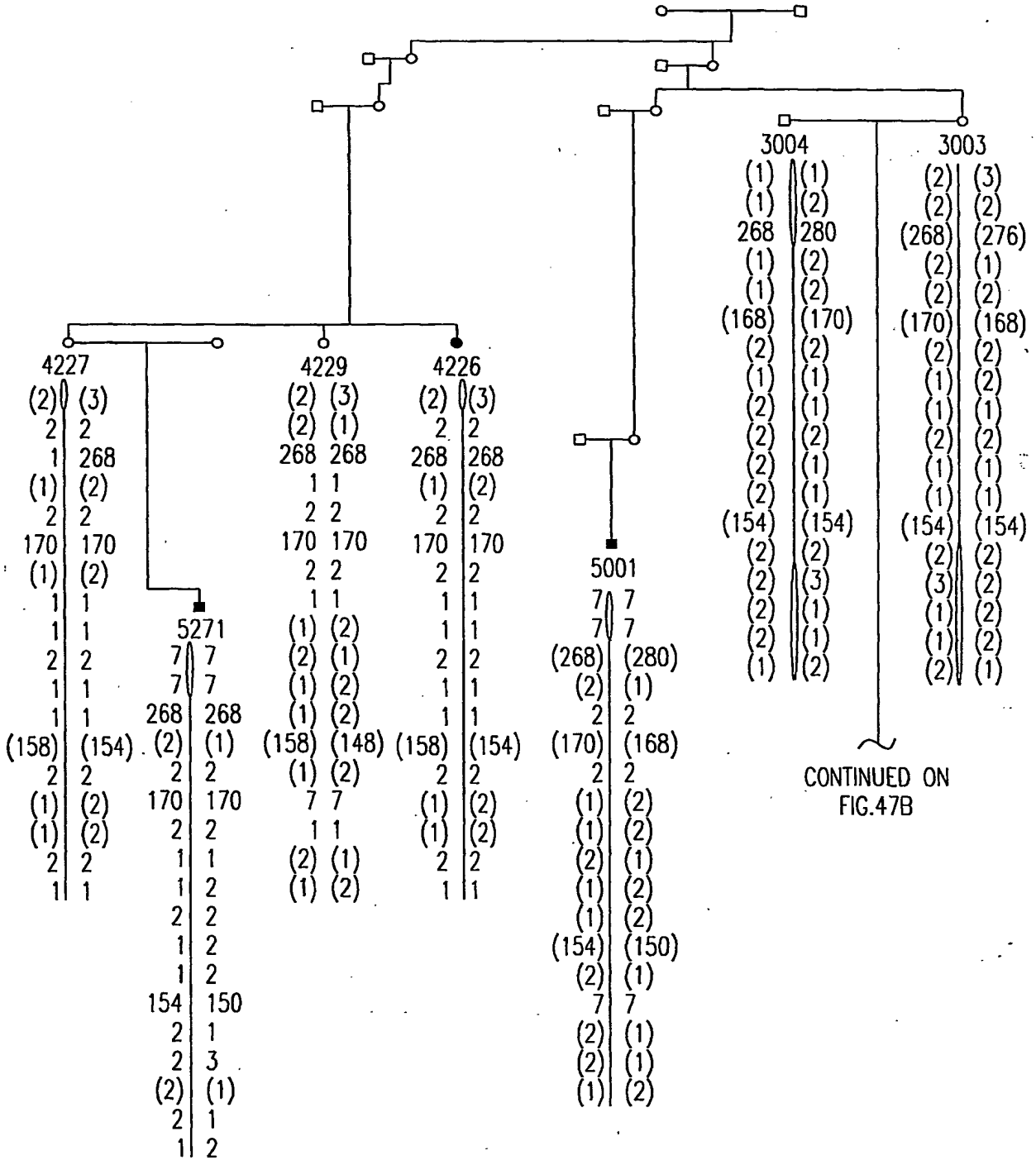


FIG.47A

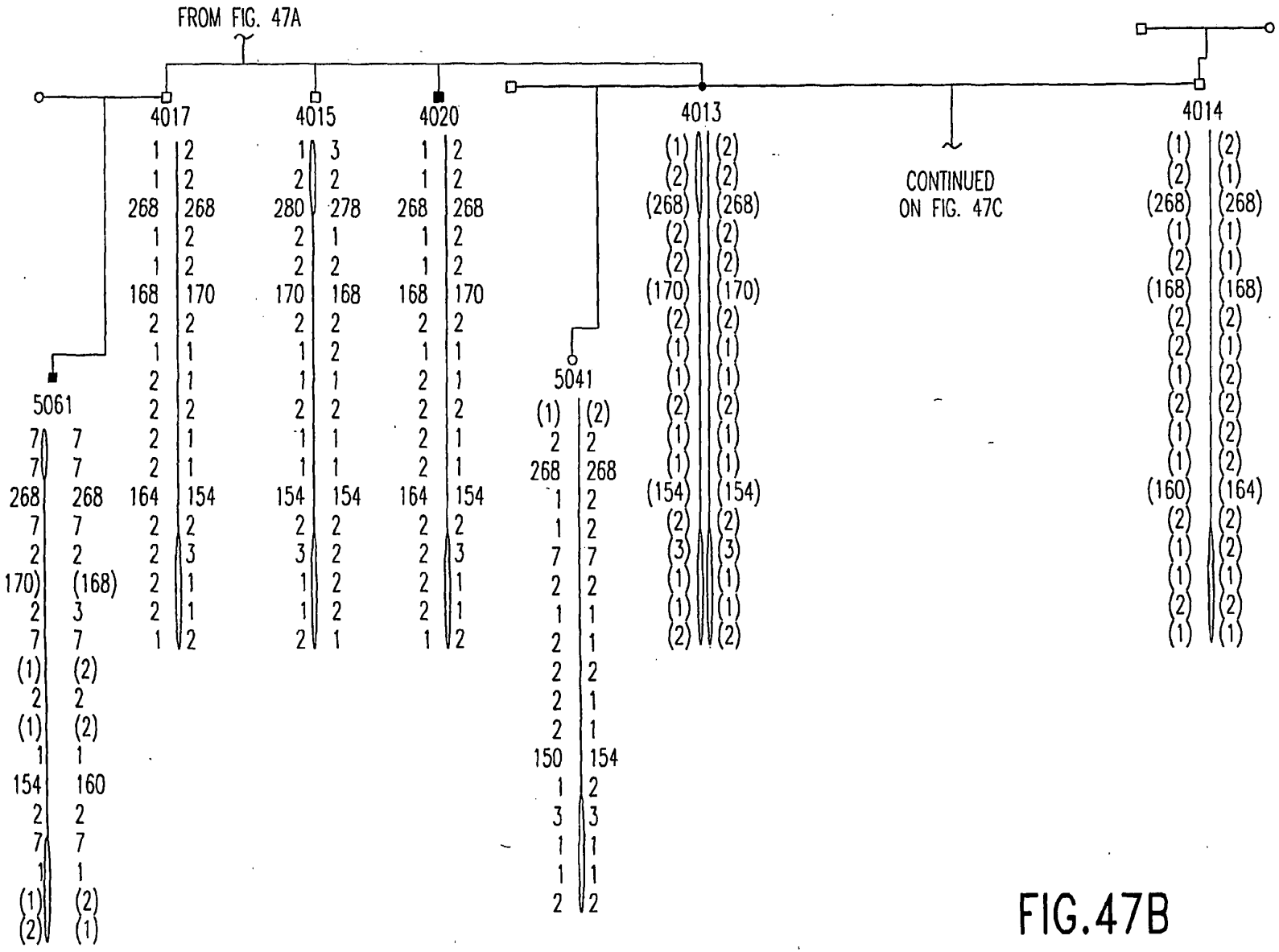

FROM FIG. 47A
4017
4015
4020
4013
4014
5061
5041
CONTINUED ON FIG. 47C

FIG.47B

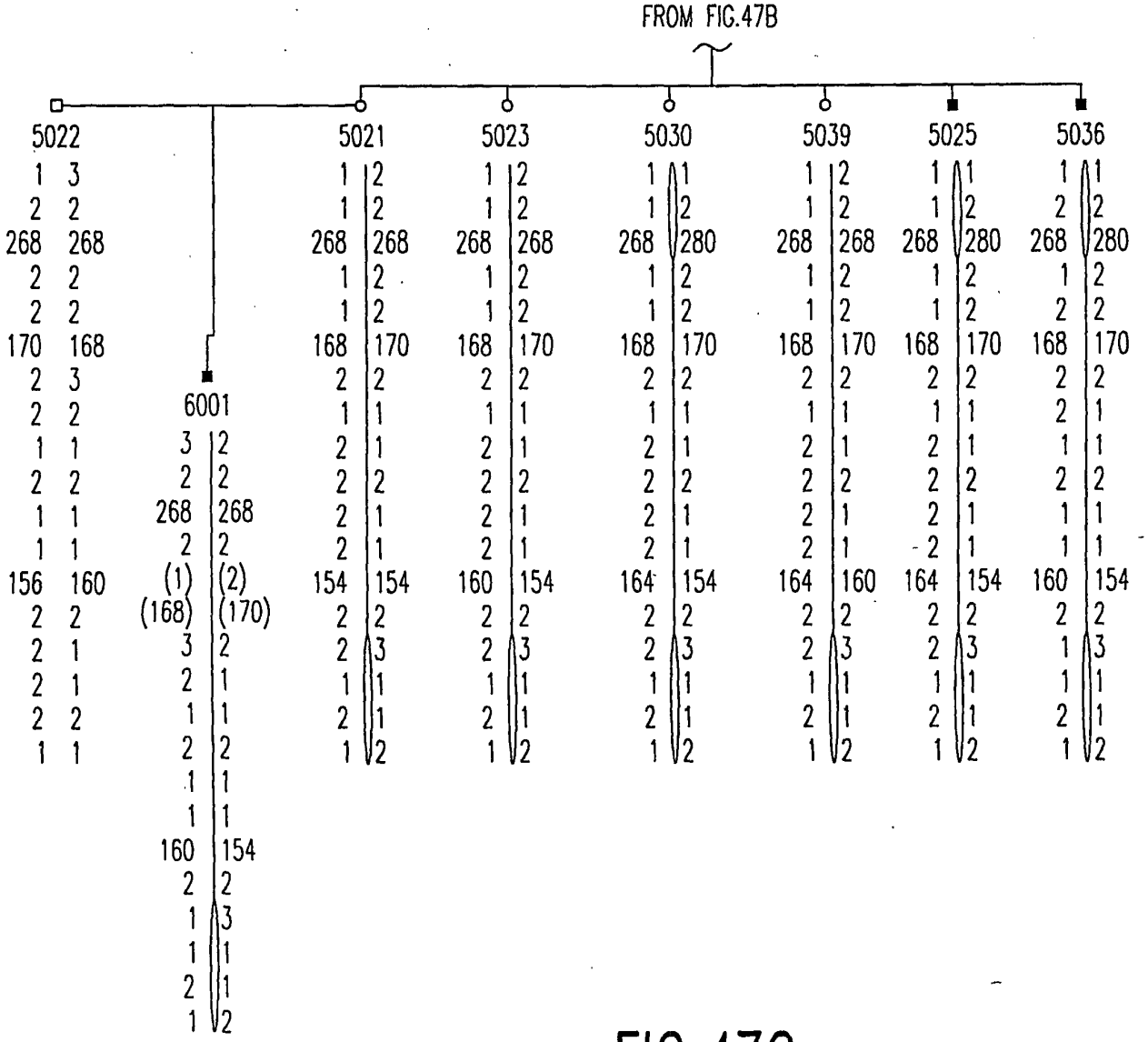

FROM FIG.47B
5022
1 3
2 2
268 268
2 2
2 2
170 168
2 3
2 2
1 1
2 2
1 1
1 1
156 160
2 2
2 1
2 1
2 2
1 1
6001
3 2
2 2
268 268
2 2
(1) (2)
(168) (170)
3 2
2 1
1 1
2 2
1 1
1 1
160 154
2 2
1 3
1 1
2 1
1 2
5021
1 2
1 2
268 268
1 2
1 2
168 170
2 2
1 1
2 1
2 2
2 1
2 1
154 154
2 2
2 3
1 1
2 1
1 2
5023
1 2
1 2
268 268
1 2
1 2
168 170
2 2
1 1
2 1
2 2
2 1
2 1
160 154
2 2
2 3
1 1
2 1
1 2
5030
1 1
1 2
268 280
1 2
1 2
168 170
2 2
1 1
2 1
2 2
2 1
2 1
164 154
2 2
2 3
1 1
2 1
1 2
5039
1 2
1 2
268 268
1 2
1 2
168 170
2 2
1 1
2 1
2 2
2 1
2 1
164 160
2 2
2 3
1 1
2 1
1 2
5025
1 1
1 2
268 280
1 2
1 2
168 170
2 2
1 1
2 1
2 2
2 1
2 1
164 154
2 2
2 3
1 1
2 1
1 2
5036
1 1
2 2
268 280
1 2
2 2
168 170
2 2
2 1
1 1
2 2
1 1
1 1
160 154
2 2
1 3
1 1
2 1
1 2

FIG.47C

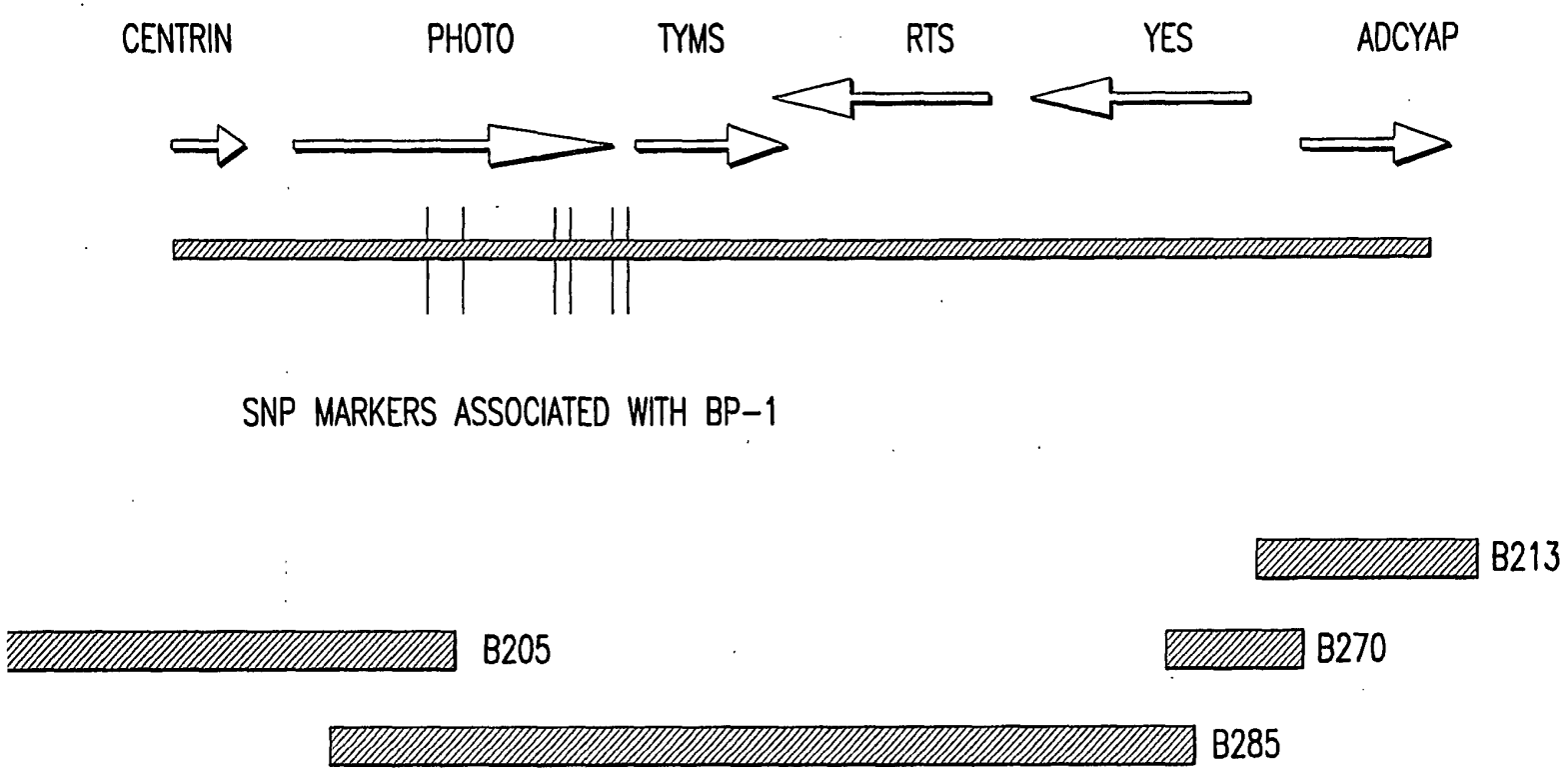

CENTRIN
PHOTO
TYMS
RTS
YES
ADCYAP
SNP MARKERS ASSOCIATED WITH BP-1
B213
B205
B270
B285

FIG.48